DAMAGE IDENTIFICATION IN KEVLAR/EPOXY COMPOSITES THROUGH CLUSTER ANALYSIS OF ACOUSTIC EMISSION DATA

by

Lt. D. RAVI KUMAR

DEPARTMENT OF MECHANICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

FEBRUARY, 1989

DAMAGE IDENTIFICATION IN KEVLAR/EPOXY COMPOSITES THROUGH CLUSTER ANALYSIS OF ACOUSTIC EMISSION DATA

A Thesis Submitted
in Partial Fulfilment of the Requirements
for the Degree of

MASTER OF TECHNOLOGY

by

Lt. D. RAVI KUMAR

to the

DEPARTMENT OF MECHANICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

FEBRUARY, 1989

CERTIFICATE

This is to certify that the thesis entitled, DAMAGE IDENTIFICATION IN KEVLAR/EPOXY COMPOSITES THROUGH CLUSTER ANALYSIS OF ACOUSTIC EMISSION DATA', by Lt D Ravi kumar is a record of the work carried out under our supervision and has not been submitted elsewhere for a degree

NN Kishore

(N N Kishore)
Asst Professor
Mechanical Engineering Department
Indian Institute of Technology
Kanpur - 208016

BD Agarwal

(B D Agarwal)
Professor
Mechanical Engineering Departmen
Indian Institute of Technology
Kanpur - 208016

Feb, 1989

ACKNOWLEDGMENTS

I take this opportunity to put on record, my deep sense of gratitude to my thesis supervisors Prof B D Agarwal and Dr N N Kishore for their valuable guidance, constant inspiration and encouragement through out the present course of work

I am deeply grateful to Mr Gopalkrishnan, Research Engineeer, ACES and Mr Nirmal Roberts, Software Engineer, Computer Centre for their valuable suggestions and timely help during interfacing

I wish to thank all the staff members of Experimental Stress Analysis Lab for their cooperation at all times Special thanks are due to Mr D K Sarkar, Mr B D Pandey, Mr Sudhir and Mr Rakesh for their dedicated help at a crucial time Thanks are due to Mr B K Jain of ACMS for his excellent drfating work

I wish to recall the efforts of my parents in bringing me up to this level A record of appreciation is due to my wife Saru for her constant encouragement and cooperation and daughter Dipika for keeping a cheerful environment at home

I would like to recall my pleasant association with my friends Mr Bajpai, Mazumdar, Sharad Gupta, Arun, Chakrovarthy, RK, Annie, Camy, Ramesh, BG Reddy, Venkat, Nasy, Ravindra and others who have made my stay a memorable one

Finally, I wish to thank all those, whose association helped me directly or indirectly in accomplishing the present work and during the course of my program at IIT Kanpur

- Lt D Ravi Kumar

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 GENERAL INTRODUCTION

When two or more distinct constituent materials or phases are combined on a macroscopic scale;they give rise to a composite material. These consist of one or more discontinuous phases embedded in a continuous phase. The discontinuous phase is called the reinforcement and the continuous phase is termed matrix. Composite materials continue to replace the conventional materials in an increasing number of a applications owing to their high strength and stiffness, coupled with low density. Today, fiber composites have found such diverse applications as space vehicles, aircrafts, offshore structures, automobiles, sporting goods, electronics etc. The use of composite materials in vital structural applications calls for a reliable testing method for their characterization and damage monitoring in these materials.

Non Destructive Testing (NDT) detects discontinuities in materials. But none of the commercially established NDT techniques can judge the injurious nature of a defect. All efforts in improving the present NDT techniques are towards extracting more information about the flaw so that the design engineer can decide about the acceptability of the defect. Also,none of the presently used NDT techniques offer real time information regarding flaw initiation or growth. Acoustic Emission (AE) technique promises to fill up the above gap,being

the most powerful tool yet known in the world of NDT technology.

Acoustic Emission is the elastic energy that is spontaneously released by materials when they undergo deformation. Monitoring of these AE signals and their analysis has proved to be an effective testing method for tracking of damage progression in a material.

When a defect such as a crack extends within a material,it generates stress waves and these are detected by Piezo-Electric transducers attached to the surface of the structure/specimen. The AE signals produced are amplified and processed before being displayed or recorded. Later, the results could be interpreted in terms of rate of crack growth,the quality of the material,or the imminence of failure of the structure depending on the location and nature of the defect.

The application of mechanical or thermal stress to a composite material results in elastic energy being stored in the material. The sudden local release of such energy is the basic cause of AE. The specific mechanisms that produce AE events in composite materials are numerous and include fiber breaks,fiber matrix interfacial debonding,matrix cracking and splitting,inter laminar debonding,rubbing of fiber against the matrix. Composites are generally noisy. Typically,the AE from a fiber reinforced composite material is of significantly larger amplitude than that from a metal. This property of fiber reinforced composites make AE testing an effective method for detection of defects as they arise,in a composite specimen during a test or a composite structure during its natural loading.

1.2 STATE OF THE ART

AE and microseismic activity are naturally occurring phenomena. The first AE detected, may well have been in making pottery which dates back to 6,500 B.C. But it was only around the turn of the twentieth century that metal researchers have started to think on AE in metals. In 1916, Czochralski [1] was the first to report, the association between "Zinn-und-Zinkgeschrei" (tin and zinc cry) and twinning.

The transition from the incidental observation of audible tin cry to the deliberate study of AE phenomena consisted of three separate and unrelated experiments. The first of them was instrumented specifically to detect AE was conducted in Germany and published in 1936 by Forster and Scheil [2]. They recorded the noises caused by formation of martensite in 29% nickel steel. In 1948, the second instrumented AE experiment was performed and published by Shockley et al [3] in United States which was directed towards observation of moving dislocations in pure tin specimens by means of the stress waves they generated. The third experiment was performed in England by Millard [4] associated with twinning on single crystal wires of cadmium.

The genesis of today's technology in AE was the work of Joseph Kaiser [5]. In 1950, he reported the first comprehensive investigation into the performance of AE. He used tensile tests of conventional engineering materials to determine: (1) what noises are generated from within the specimen; (2) the acoustic process involved; (3) the frequency levels found; and (4) the

relation between the stress-strain curve and the frequencies noted for the various stresses to which the specimens were subjected. He concluded that the occurrence of AE arises from frictional rubbing of grains against each other and also form intergranular fracture. His most significant discovery was the irreversibility phenomenon which now bears his name, the Kaiser effect. Kaiser effect can be defined as the absence of detectable AE until the previous maximum applied stress level has been exceeded.

In 1969, a new approach to NDT was proposed by Dunegan and Harris [6] which was centered around the concept of utilizing the transducer action of a flaw in a stress field. In 1971, Pollock [7] directed his research towards the practical application of AE for material testing and NDT.

The application of this testing technique to the field of composite materials was suggested by Anderson and DeLacy [8]. In 1974, Hamstad [9] and Carlyle [10] examined the research and development applications that AE monitoring has for fiber reinforced composites. While Hamstad's work was a general and comprehensive guide to the use of AE testing, Carlyle placed specific emphasis upon obtaining information from fundamental failure process in composites. Guild [11] also examined the state of the art with regard to the composite materials.

Arrington [12] used AE for debond tests in single fiber specimens. In these tests, single fibers were embedded in specially shaped resin specimens which were tested in compression resulting in debonding. The AEs observed were directly related to debonding. Thereby, the load at the initiation of debonding and

the severity of debonding as the load increased, could be determined. This was then used to obtain the bond strength.

Ahlborn et al. [13] used AE as a indicator for the beginning of crack propagation in notched specimens of unidirectional reinforced epoxy-resin during bending and tension tests. They observed AEs even when the areas of delamination or debonding were smaller than 1 mm^2 and thus concluded that AE can be used effectively to indicate the beginning of crack propagation. Bunsell et al.[14] worked on evaluation of AE as a means of monitoring damage within composite materials.

Brown [15] investigated AE from specimens of glass reinforced plastics. His emphasis was on the use of AE for material assessment rather than for defect location. He conducted tensile tests on different specimens classified in terms of matrix type - epoxy or polyester, resin hardener ratio and density of square woven-roving reinforcement. He used ring down counts (RDC) as the parameter for his analysis. Conclusions took the form of empirical relationships based on the analysis of adequate number of identical specimens.

Chiao et al. [16] investigated the performance of Kevlar-49 with regard to fiber uniformity, strength distribution at room temperature, stress-strain characteristics, strain rate effect on fiber strength, stress-rupture behaviour, and aging under no load. He also carried out microstructural characterization of the failure mode of the fiber using electron and optical microscopy. Hamstad and Patterson [17] studied AE monitoring of spherical Kevlar epoxy composite pressure vessels. His study was focused on

three specific areas:1) development of an experimental technique and proper instrumentation to measure the energy given off by the AE transducer per AE burst;2) design of a test fixture for mounting the composite vessels; and 3)the number,location,and sensitivity of AE transducer used for proof testing of composite pressure vessels.

Right from the time when the AEs were first observed till today, investigators have been constantly conducting research to exploit the AE technique in the most exhaustive fashion. In 1983 Bunsell [18] provided a means of following the accumulation of damage in carbon fiber reinforced resin structures. He has suggested amplitude analysis of the emissions as the identification of the sources of emissions presented some difficulties. Two years later, Moore et al. [19] carried out tests and monitored AE during single filament tension tests on Kevlar-49 fibers. They studied the behaviour of dry and wet fiber bundles and conducted statistical studies on the AE parameters i.e., peak amplitude, energy, ringdown counts and event duration. They concluded that energy changed more than event duration,counts or peak amplitude did from their typical values for a single filament break in the bundle.

Awerbuch and Ghaffari [20-22] directed their study towards discriminating friction generated emission from those caused by actual damage progression through a proper correlation among the AE event intensity variables. While working with graphite-epoxy specimens subjected to fatigue loading [21], quasi-static loading [22] and impact damaged composites [20], it was demonstrated that a significant amount of emission is generated by fretting among

the fracture surfaces developed during loading.

Clerico et al. [23] in their paper aimed at evaluating the conventional NDT technique,utilized for composites (AE and passive AE) in detecting porosity,simulated delaminations and slight impact damage on already used and new epoxy and bismaleiymide thermosetting composites. These were prepared with intermediate modulus carbon fibers for improving the toughness characteristics. They concluded that AE monitoring could be a powerful tool for obtaining information about the integrity of a structure,especially impact damage process.

Cohen and Awerbuch [24,25] studied the applicability of AE for detecting initiation,accumulation,and progression of delamination in graphite/epoxy laminates. During the fatigue loading [24] and quasi-static loading tests [25],it was observed that a significant amount of emission is generated by the extensive grating among the fracture surfaces. This exceeded the emission that was caused by the actual progression of the delamination by several orders of magnitude,depending on the duration of fatigue loading. They pointed that either to establish a correspondence between the modes of failure and AE event intensities or to monitor damage accumulation and track its progression, the friction emission should be identified and separated. For this purpose, experiments were performed with a model composite in which delamination is the single dominant mode of damage. It has been shown that when the friction emission is identified and separated,a better correlation can be established between the occurrence of delamination and the amplitude distribution histograms of the events and also between the

extension of the delamination, as determined from AE and the actual progression measured from optical observations.

In 1988 Tasnom et al. [26] developed a comprehensive AE monitoring method for composite materials with regard to two fundamental problems:

i) To distinguish between failure modes in composites such as fiber failure, delamination, matrix micro-cracking, etc.

ii) How to evaluate the contribution of individual and combined failure modes to the overall damage process and ultimately to the actual residual strength and life of a composite.

Selected AE experiments with a carbon-epoxy composite made of high strain carbon fibers (Hercules IM6) and a Bismalimide modified epoxy (Narmco 5245C) were carried out to demonstrate the fundamental basis of the comprehensive AE monitoring approach. He concluded that AE digital signal analysis offers much potential to characterize the source mechanisms by frequency analysis or by deconvolution.

One of the recent investigations were carried out by Awerbuch and Bakuckas [27] on metal-matrix composites (MMC). Damage initiation and accumulation was monitored through Ae in several center-notched unidirectional MMCs and multi- directional boron/aluminum laminates subjected to uniaxial quasi-static tensile loading. It was shown that the failure process in the different material systems could be correlated with AE event amplitudes and also that an excellent correlation could be

established between the rate of damage growth and AE results. Thus, it was concluded that AE technique can monitor damage formation and accumulation in MMC in real time.

Several investigators made use of statistical analysis to characterize and identify damage mechanisms in composite materials. Williams Jr et al.[28] used statistical analysis procedure to distinguish the different predominant failure mechanisms in graphite epoxy specimens. Groups of AE events were treated as random data and statistically analysed to identify group characteristic which enable mechanism discrimination. Prasanna Kumar et al. [29] used AE a material testing tool to test aluminum base alloys through amplitude distribution analysis,statistical analysis and time series analysis. Philips and Harris [30] on the basis of amplitude distribution obtained from several types of glass laminates concluded that the AE source mechanisms cannot be determined on the basis of amplitude distribution alone.

Abdul Majid [31] applied unsupervised pattern classification technique to investigate source characterization of AE signals. He observed that by using the threshold K-means method the sources can be characterized and feasibility for pattern recognition exist.

From the summary of the literature survey it can be inferred that the ultimate goal of various investigators irrespective of the approach,was to detect/monitor damage initiation and accumulation. They then made in-depth study to determine the damage criticality with regard to AE event parameters and specific failure process such as matrix cracking and

splitting, fiber breaking, debonding, delamination and fiber pull-out.

1.3 PRESENT WORK

The present work is an attempt to develop a comprehensive procedure to identify damage mechanisms in Kevlar/epoxy composites through cluster analysis of Ae data. AE data was recorded in real time by interfacing 'AET 5000 Microcomputer based AE monitoring System' to MTS-810 system. Northstar Advantage is the microcomputer that was used with AET 5000 which has a memory of 64K.

A typical AE system consists of sensors, preamplifiers, band-pass filters, main amplifier, processing unit, graphic display and recorder. Details concerning principles and working of AE system and the explanation of various AE terms are contained in chapter 2.

The present work relates to analysis of AE data obtained for Kevlar fabric reinforced epoxy resin laminates of $[0_{12}]$, $[90_{12}]$, $[45_{12}]$, $[0_3/90_3]_s$, $[90_3/0_3]_s$ configuration and plates made of fiber rich laminates. Cluster analysis programs demand huge memory storage and processing time. Keeping this demand in mind, the Northstar Advantage was interfaced with "Dynalog PC/XT". A brief general experimental procedure and data acquisition from AET 5000 along with the interfacing details are described in chapter 3.

Cluster analysis has been employed as an effective tool in scientific inquiry and in a great variety of applications. Introduction to cluster analysis, types of methods, details of

software developed for analysis and post processing are presented in chapter 4.

The results obtained from the software developed are discussed in chapter 5. Conclusions drawn and some suggestions for future work have been included in chapter 6.

CHAPTER 2

PRINCIPLES OF ACOUSTIC EMISSION

2.1 THE AE PHENOMENON

An AE, as mentioned earlier is the class of phenomena where transient elastic waves are generated by the rapid release of energy from localized sources within a material. These transient elastic waves are sometimes referred as pressure or stress waves. The acoustic energy, is usually in the form of short bursts or trains of fast impulses. These internal releases of energy are detected by AE testing by using transducers to pick up the transient elastic waves. These signals can then be related to the physical integrity of the material or structure in which they are generated. Monitoring of these events permits detection and location of flaws as well as prediction of impending failure. Fig 2.1 represents generation of a typical AE signal.

The AE signals may be broadly divided into two types: continuous and burst type. continuous emission signals may originate from such sources as leaks in pressurized systems, yielding in metals, hydraulic noise and rotating machinery noise. The method most accepted for measuring continuous emission is by rms or "signal level" technique, as displayed in fig 2.2. This method does not use a threshold voltage, thereby including, even small variations from the background noise in the measurement.

Burst-type emissions have the characteristic AE waveform displayed in fig 2.3. These signals originate from a variety of sources, including crack growth in metals, composites, geologic materials, fluid cavitation, metal fretting and material impact. These signals are usually characterized by a fast rise time to

peak, followed by an exponential decay. The most commonly reported AE from fiber reinforced composites is burst emission.

Sources of AE include many different mechanisms of deformation and fracture. These can be basically classified into four different groups: (1) Dislocation movements, (2) Phase transformations, (3) Friction mechanisms and (4) Crack formation and extension. The sources of specific mechanisms that produce AE in composite materials are numerous and include the following: matrix cracking, fiber-matrix debonding, fiber break, fiber pull-out and delamination. These mechanisms typify the classical response of materials to applied load. Typical damage mechanisms in fiber reinforced composites are shown in fig 2.4.

2.2 ADVANTAGES AND DISADVANTAGES OF AE TESTING

ADVANTAGES :

1. Real time evaluation
2. High sensitivity (discontinuities as small as about 10^{-11} mm in metals, and fractures of single filaments of organic fibers with diameter of 10^{-2} mm can be detected)
3. Total specimen volume sensitivity
4. Clear location of damage regions
5. Continuous monitoring
6. Pattern recognition aids interpretation

DISADVANTAGES :

1. Difficulty in discrimination between noise and AE signals
2. Material behaviour must be understood
3. Quantitative correlation is limited
4. Requires experience

2.3 AE SYSTEM :

A typical AE system consist of sensors, pre-amplifier, main amplifier, filters, measurement circuitry, data buffers, micro computer, graphic display and recorder. The AE testing and flaw detection involves conversion of AE to electrical signals, their amplification, filtration, processing and recording. Block diagram of a general AE system is shown in fig 2.5.

The test equipment used is AET 5000, a computer based general purpose AE monitoring system comprising a computer automation LSI4/10 microprocessor, a graphic display terminal (GT) and a Epson FX80 hardcopy unit. Detection of AE is accomplished by the sensor acting through a couplant. The couplant is used to fill the air gap between the sensor shoe and the surface of the specimen to ensure good transmission of AE energy from the sample to the sensor. The piezo-electric sensors are attached to the specimen through a couplant SC6. They convert stress waves traveling through a medium into minute variations of emf (the AE signals).

The strength of AE signals are normally very low and require more than 60 dB amplification, which is obtained through a pre-amplifier. This high gain requirement is due to the relatively low amplitudes of AE signals from composites, when compared to metals. The preamplified signal is passed through a bandpass filter and a main amplifier, with the objective to optimize the signal to noise ratio. The function of filters is two fold: to eliminate the background noise and to transmit the signal in the desired frequency range. The conditioned (amplified and filtered)

AE signal is compared to an internally generated reference voltage (the "threshold"). The comparator (or signal processing unit, SPU) emits a pulse each time the AE signal crosses the reference voltage. Each pulse from the comparator represents a ring down count. The SPU outputs the digital AE signal to the AE modules (TDM, REM, ARM AND PRM-1).

The time difference module (TDM) measures the difference in time of arrival of an AE wave at two different sensors. Thus one TDM is required for the linear location determination of AE sources whereas two TDM's are required for planer location. Sensor locations are assigned through keyboard entries.

The ringdown counter/event duration module (REM) counts the number of threshold crossings in an AE event, counts the number of clock pulses to determine the event duration and supplies a control to the amplitude/rise time module (ARM).

The ARM receives the AE signal directly from the pre-amplifier output. This signal undergoes logarithmic amplification, and is converted to a voltage representing dB. It measures the peak amplitude (in dB) of an event referenced to the fixed gain (60 dB) of the pre-amplifier. The rise time clock registers the time from first threshold crossing to the occurrence of peak amplitude. The processor operates on the peak amplitude/rise time data, determining the positive AE signal slope. The parametric/RMS module #1 (PRM1) is an analog-to-digital converter (ADC) which measures and digitizes the system power voltages, the signal level voltages and three user supplied external analog voltages that may be associated with the AE test. These external voltages may represent such test conditions as

load, displacement, pressure, temperature etc. The signal is then processed by the AET 5000 system and data recorded to disk on the Northstar Advantage intelligent terminal via a parallel interface from the LSI4/10 main system computer.

The Advantage becomes a user interacvtive terminal for the AET 5000 by running the program AET.COM, that causes the Northstar to emulate a Tektronix graphics terminal. In addition to the graphics capability, AET.COM also store test data,plays it back for processing by the AET 5000. The GT has a CRT terminal with both text and graphic memories. The GT data displays can be hardcopied on the Epson FX80 dot matrix printer connected to the GT through a composite video interface.

2.4 GLOSSARY OF AE TERMS

Acoustic Emission Signal : The signal obtained by the detection of AE is known as acoustic emission signal.

Threshold Crossing : Whenever an AE signal voltage exceeds a preset referance voltage (i.e. the threshold) a threshold crossing is said to have occurred.

Acoustic Emission Event : An acoustic emission event is the AE generated as a result of a local material change. An event is said to have begun when the AE signal amplitude exceeds the preset threshold for the first time and is said to be over when the next threshol crossing is not observed within the specified time gap (time gap is set so that if the gap between the two threshold crossings exceeds this value, the former threshold crossing is considered as the last of the previous event while the later becomes the first of the new event).

Ring Down Counts : The number of times the AE signal amplitude exceed a preset threshold during an AE event is known as the ring down counts (RDC) of that event.

Event Duration : Event duration (ED) is the time measured from the first threshold crossing to the last threshold crossing of an event. It is measured in micro seconds.

Peak Amplitude : Peak amplitude (PA) is the measure of the peak signal in an AE event. It is measured in decibels (dB).

Rise Time : Rise time is the time taken for an AE event to crosses the first threshold crossing to the occurrence of the peak amplitude. It is measured in micro seconds.

Slope : Slope is the measure, as to how fast the peak amplitude is observed in an AE event. It is a software calculated parameter obtained as peak amplitude divided by rise time.

Energy: Energy, when referred in this study, is meant by the acoustic energy released by an AE event. This is a parameter calculated as:

$$Energy = PA + 10 \, Log_{10} \, ED$$

Arrival Time Difference: Arrival time difference (DT) is the time interval between the detected arrivals of an AE wave at two different sensors of a sensor array.

Acoustic Emission Intensity : Any of the several measurable qualities of AE signals from an active acoustic source that may be used to grade the severity of the source can be called as acoustic emission intensity. Qualities that are measures of intensity are RDC,PA,ED,RT,Energy etc.

Decibel and Gain : Decibel is the log of a ratio. Gain is

the measure of signal amplification. For a ratio of voltages we have:

$$\text{Gain (dB)} = 20 \log_{10} V_2/V_1$$

Thus, an amplifier that produces one volt of output (V_2) for a one millivolt input (V_1) has a gain of 60 dB.

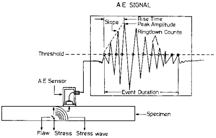


Fig. 2.1 AE Signal Generation

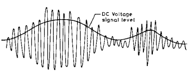


Fig. 2.2 Signal level (rms) method of measurement of continuous AE signals.

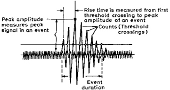
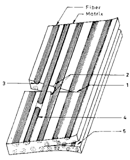


Fig. 2.3 Characterization of burst emissions by the thershold method.



1 Matrix Cracking
2 Fiber - Matrix Debonding
3 Fiber Break
4 Fiber Pull out
5 Delamination

Fig. 2.4 Sources of AE in fiber reinforced composites

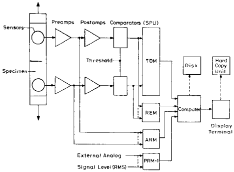


Fig. 2.5 Block diagram of a general acoustic emission system.

CHAPTER 3

EXPERIMENTAL PROCEDURE

3.1 INTRODUCTION

This chapter describes cluster analysis of AE data obtained for Kevlar fabric reinforced epoxy resin laminates of $[0_{12}]$, $[90_{12}]$, $[45_{12}]$, $[0_3/90_3]_s$, $[90_3/0_3]_s$ configuration and fiber rich laminates. The specifications of Kevlar-49 and epoxy resin as supplied by manufacturers are given in Table 3.1 and Table 3.2 respectively. As mentioned earlier, Northstar Advantage, the microcomputer of AET System 5000, was interfaced with "Dynalog-PC/XT", primarily for dat transfer. A brief experimental procedure, from specimen preparation to data acquisition from AET 5000, along with the interfacing details are described in the following sections.

3.2 COMPOSITE LAMINATE FABRICATION

Material fabrication involves casting of composite laminate plates prior to preparation of specimens. Twelve layers of fabric was cut to proper size by a special pair of scissors intended for Kevlar cutting. The fabric was demoisturised by heating it in an oven at 105°C for 16 hrs. and cooled in the oven itself to avoid any moisture regain prior to processing. About 450 gms of epoxy resin (araldite LY556) was preheated to about 100°C and the temperature maintained for 2 hrs. to remove absorbed moisture. The resin was then allowed to cool to room temperature.

Composite laminates were cast by hand lay-up technique between two 25 mm thick M.S. plates. The plates were nickel coated at the top and affixed with a heating element at the

bottom. A mylar sheet was placed on the lower mould plate. The resin and hardener were thoroughly mixed. The amount of hardener was 10% of epoxy used. A layer of resin was spread over the mylar sheet and the first Kevlar fabric was placed. Epoxy was applied over the fabric. This process was repeated till all the twelve layers of fibers were placed. To ensure proper fiber impregnation an aluminium roller was applied after the third layer. Another mylar sheet was placed over the top most layer. Steel spacers of 3 mm thickness were placed at all four corners between the two mylar sheets. A rubber roller was rolled over the top sheet to squeeze out excess epoxy and entrapped air. The upper mould plate was then placed in position. The mould plates were secured by uniform tightening of nuts on the bolts, provided at the four corners. Fig 3.5 shows the entire set up. The laminates thus prepared were cured for 6 hours at room temperature followed by 16 hours curing at 55-60°C. After curing, the composite plate along with mylar sheet was removed. The mylar sheets were then removed to get a fine finished composite plate.

The above technique was used in making laminate of configuration $[0_{12}]$, $[45_{12}]$, $[90_{12}]$, $[0_3/90_3]_s$ and $[90_3/0_3]_s$. Volume fraction of fibers (V_f) was found to be 52.8% using the following formula (since Kevlar fibers defy the standard resin burn-off test) :

$$V_f = (AN \times P_{fa}/P_f) / (A \times t) = (N \times P_{fa})/(t \times P_{fa})$$

where

A : Area of the composite laminate

TABLE 3.1 : KEVLAR-49 FABRIC SPECIFICATIONS

Product	-	Du Pont Co.,USA
C.S. Style	-	343
Former Du Pont Style	-	143
Weight (per unit area of fabric)	-	190 g/m^2
Tensile Strength Warp	-	255700 N/m
Fill	-	25700 N/m
Count (number of yarn per inch in warp x fill)	-	100 x 20
Yarn Denier (weight in grams for 30,000 feet long yarn) warp	-	380
fill	-	195
Weave	-	Crowfoot
Finish	-	CS - 805
Fiber Properties:		
Specific Gravity	-	1.44
Decomposition Temperature	-	500°C

TABLE 3.3 : EPOXY SPECIFICATION

Product	-	CIBA Geigy India Ltd.
Category: Resin	-	Araldite LY556
Hardener	-	Hardener HY951 (10% of araldite by weight)
Viscosity	-	5000 - 8000 cp
Pot Life	-	30 minutes to 1 hr.
Specific Gravity	-	1.2 - 1.3
Tensile Strength	-	55 - 130 Mpa
Tensile Modulus	-	2800 - 4200 Mpa
Poisson's Ratio	-	0.2 - 0.33
Flexural Strength	-	125 Mpa
Decomposition Temperature	-	270 - 280 °C

N : Number of fabric layers in the laminate

P_{fa} : Area density of the fabric

P_f : Density of fiber

t : Thickness of the laminate

In the present case we have,

N : 12

P_{fa} : $0.190\ Kg/m^2$

P_f : $1.44 \times 10^{-3}\ Kg/m^3$

t : 3×10^{-3}

3.3 FIBER RICH LAMINATE FABRICATION

Twelve layers of Kevlar cloth fabric of size (30 x 25 mm) was weighed after demoisturization in the oven for 15 Hrs. Epoxy (10% by weight of fiber) was prepared. The laminate was then prepared exactly as in hand lay-up technique. These laminates were prepared to find the AE characteristics of fiber. The small percentage of epoxy used in the laminate preparation was to just bind the layers so that, signal transmission is possible. The thickness of laminate plate thus obtained was 1.5 + 0.01 mm. The volume fraction of fibers were 80 to 85%.

3.4 PREPARATION OF COMPOSITE AND FIBER RICH SPECIMENS

Specimen preparation involves cutting the specimens of desired size from the laminates, finishing of specimens and cutting the notches in these specimens.

Specimens of the dimensions, as shown in fig 3.6 were cut from the fabricated laminates. Kevlar composites, due to their high toughness and strength pose some difficulties during mechanical cutting. The common problems that are encountered

are : at low cutting speeds, instead of cutting at the edge, fibers break and pull out from inside of the composite and brooming at high speeds.

The best results in cutting laminates were obtained in circular sawing when the unconventional side (slant side of teeth) of a metal sliting fine teethed HSS cutter was used at high speeds. Water was used as coolant. With this set up, cutting speed in the dense fiber direction (cutting less fibers) was quite high and in cross direction it was moderate. Brooming was very little except in crossply laminate at the edge, on the side from which the cutter comes out after cutting. The broomed fibers were removed by a special pair of scissors and the edge was finished by using water proof emery paper. During finishing operation it was ensured that the final dimensions of the specimen are as par the requirements.

The specimens were then marked and single edge notches were cut on a lathe using 0.5 mm thick slit cutter and the a/w ratio (notch length to specimen width ratio) was kept at 0.6. All the three operations cutting,finishing and notch making were done with due care and the coolant was used in each of these operations to ensure a proper finish.

3.5 PRELIMINARY SETTINGS ON AET 5000 SYSTEM

During the performance of AE tests various settings were fed through keyboard entries as listed below:

The distance between the two sensors, used in these tests, was divided into 100 segments for the purpose of monitoring the line location. These segments correspond to location "0" to "100"

between the sensors.

The preamplifier gain for both the sensors were set to a value of 60 dB. The threshold level was set at a value of 1 volt fixed. The machine noise through grips and other noise levels were observed to be less than this value after mounting the specimen. Type of tests conducted was linear, with sensor numbers 1 and 2 having locations of "0" and "100" respectively. Maximum DT was initially set at 0 which gets automatically corrected to a value obtained during the course of calibration.

Clock period of event duration module was set at 125 ns and that of rise time module as 250 ns. Ranges of different AE parameters that were used during test are shown in table 3.3.

TABLE 3.3 : RANGES OF VARIOUS AE PARAMETERS

	MIN	MAX
Event Duration (μs)	0	65520
Ring Down Counts	0	4096
Rise Time (μs)	0	65520
Peak Amplitude (dB)	0	117
Slope	0	65520
Energy	0	165
Analog Parameters	0	10240

3.6 EXPERIMENTAL PROCEDURE

Basic characterization of the material was done by tensile

tests of unnotched specimens of $[0_{12}]$, $[90_{12}]$ and $[45_{12}]$ configuration on a 10 tonne MTS model 810 materials testing system. The material properties obtained are given in table 3.4. The notched specimens were tested by interfacing the AET 5000 and MTS-810 system.

The specimen surface was cleaned with acetone and made dirt-free. Two sensors were then attached to the specimen surface with their centers 10 cm apart and equidistant from the central notch. The sensors were attached with the help of a couplant, SC6 silicon

TABLE 3.4 : ELASTIC CONSTANTS OF THE MATERIAL

Longitudinal Modulus	(E_L)	-	49.08 Gpa
Transverse Modulus	(E_T)	-	12.88 Gpa
Inplane Shear Modulus	(G_{LT})	-	2.65 Gpa
Major Poisson's Ratio	(ν_{LT})	-	0.193
Minor Poison's Ratio	(ν_{TL})	-	0.051

grease smeared on the shoe of the sensors in a light layer.

Firstly the calibration of the test was performed with each specimen to get a representative speed of sound in the specimen material. The calibration is to be done in case of every distinct specimen configuration since the specimen geometry has an influence on the result. Calibration was performed by placing a standard AE source in terms of another sensor which is connected to the mainframe's 5 volts output pulse, at a small distance from the sensor, outside the 100 segment region. The standard AE

source simulator sends out a signal of constant level. A constant output is more important than the actual level. The difference in arrival time (DT) of this signal at the two sensors was used by the system computer to determine the representative speed of sound in the material.

The pulser was then detached and the specimen was mounted on MTS machine. Mounting was done in load controlled mode. Hydraulically operated grips of the machine help in the speedy insertion and removal process of the specimen in addition to providing excellent axial and angular alignment. The MTS load and stroke digital indicators were set at a positive value of about 0.02 volts. This was done to avoid any overflow. The AET was then completely interfaced to MTS-810 system by connecting the analog parameters probes in the specific sequence. Load and displacements were recorded on AET 5000 system as external analog parameters 1 and 2.

The various test parameters were:

Load	10% (1000 Kgs)	i.e.	10 volts = 1000 Kgs
Stroke	10% (10 mm)	i.e.	10 volts = 10 mm
Specimen length between grips			= 15 cm
Time for test (stroke controlled mode)			= 10 minutes

The tests were conducted in stroke controlled mode so that the load displacement behaviour beyond maximum load is also clearly indicated. Data recording on AET computer was initialised by system commands and by opening a file for data collection. The AE system computer and MTS 'RUN' were operated simultaneously to begin the test. Tensile test was performed upto fracture. The end of the test was done by closing the data file properly and

issuing a 'END RUN' command. Fig 3.7 shows the overall view of the specimen under test on MTS and AET 5000 monitoring system.

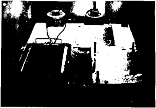
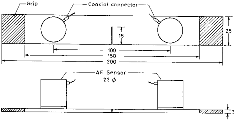


Fig. 3.5 Composite laminate fabrication.



All dimensions in mm

Fig. 3.6 Composite specimen geometry.

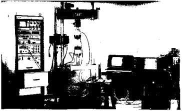


Fig. 3.7 Overall view of AE system with specimen under test on MTS.

CHAPTER 4

INTERFACING NORTHSTAR ADVANTAGE WITH PC/XT

4.1 THE NEED FOR INTERFACING

The Northstar Advantage (NS) working on CP/M operating system has 15 MB Winchester and a floppy drive. However, the drive could not be booted due to an error in the Winchester tracks. The absence of Winchester undoubtedly lowered the capability of the microcomputer in terms of speed and efficiency to a large extent. With system program, assembly language routines and graphic manager occupying 44K out of the available 64K memory, only 20K was available to the user for programming. This space is highly insufficient for handling large programs,especially cluster analysis programs.

Another major drawback is the inherent system speed. NS being run on a 8 bit processor has very slow computational rate compared to a PC/XT or DEC 1090 System. In view, of the above disadvantages it was very essential to transfer the AE data from the CP/M system after the test to an MS-DOS PC/XT, for subsequent transfer to DEC 1090 system for cluster analysis of the same. With the above concept, one primary objective was to implement a file transfer facility to enable data transfer from the CP/M system to be processed on the PC for reasons of speed and efficiency.

4.2 THE RS-232-C STANDARD

RS-232-C is defined as an interface between data terminal equipment (DTE, typically a computer or a computer terminal) and

data communication equipment (DCE, typically a modem), employing serial binary data interchange. The word modem is a contraction of modulator-demodulator. A modem is a unit incorporating a technique for placing and receiving computer signals over the common barriers. As the definition states, RS-232-C is simply a standard. This standard outlines the set of rules for exchanging data between business machines. These machines can be terminals, printers, computers or other equipment employing serial communication.

Components of a communication environment can easily be compared to a railway system. To put it in Seyer's [32] words: Just as a train is transported along the rails between stations, so is information transferred between business machines. The elements of a railway system, such as the rails, trains and stations provide an excellent illustration for comparison with the specifics of an RS-232 communication environment.

In an RS-232 environment, the departing trains are the transmitted data and the arriving trains are the received data. A RS-232 interface consists of 25 pins or leads for use by the DTE and DCE (see fig 4.1). However, only a limited number of leads are used as 'tracks' for data transfer. Each lead has a preassigned function. In general, transmitted data at the originating DTE are on pin 2 of the RS-232 interface,whereas at the receiving DTE, these same data arrive on pin 3. But in the present case, you would find that pins 2 and 3 of NS are directly connected to the corresponding pins of PC/XT. This is because, it was found that the pins 2 and 3 of NS were internally connected. Externally, though it may be seen that pins 2 and 3 are connected

directly between the computers, it is actually, pin 2 of NS being connected to pin 3 of PC/XT. i.e. AE data transmitted on pin 2 of NS arrive on pin 3 of PC/XT.

The cable that was developed for interfacing, consists of nine pins. The pins, their functions, and directions regarding (from,to) are as shown in fig 4.2. The present cable configurations were designed with respect to the cable connection between NS and EPSON FX-85 printer. Pin 1 is called protective ground which is usually, a frame ground to keep people from receiving shocks in the event of electrical shorts. Signal ground (pin 7) is used as a reference for other signals on the interface. Practically speaking, in an interface set up, similar to what was developed, essentially requires four pins (pins 1,2,3 and 7). The other pins are better understood in a communication environment.

Let us consider, control of a two way, non-simultaneous path handled by a DTE-DCE interaction as shown in fig 3.10. RTS(pin 4) of DTE 1 raises a 'Request to Send', if it has data to transmit. DCD(pin 8) is checked by modem 1 to see if DTE 2 has its RTS high (i.e.DCD is on). If the RTS is high on DTE 2, modem 1 does not give CTS to DTE 1, and subsequently DTE 1 drops its RTS signal. If DCD is off, modem 1, after a slight delay, gives a CTS (pin 5) signal to DTE 1. DTE 1 then presents data on the transmitted data lead (pin 2), and the modem 1 transmits this to the far end. Modem 2 puts the received data on pin 3 for presentation to DTE 2. DTE 1 continues with RTS held high until all data are transmitted. Then it drops its RTS, which drops DCD at the far

end and CTS locally, causing the line to idle once again. Either

Pin Number		Function
1	PG	Protective Ground
2	TD	Transmitted Data
3	RD	Received Data
4	RTS	Request to Send
5	CTS	Clear to Send
6	DSR	Data Set Ready
7	SG	Signal Ground
8	RLSD	Received line signal Detector
9	-	Reserved for data set testing
10	-	Reserved for data set testing
11		Unassigned
12	SRLSD	Secondary received line signal Detector
13	SCTS	Secondary Cleat to Send
14	STD	Secondary Transmitted Data
15	TSET	Transmission signal element timing (DCE)
16	SRD	Secondary Received Data
17	RSET	Receiver signal element timing
18		Unassigned
19	SRTS	Secondary Request to Send
20	DTR	Data Terminal Ready
21	SQD	Signal Quality Detector
22	RI	Ring Indicator
23	DSRS	Data signal rate selector (DTE/DCE)
24	TSET	Transmit signal element timing (DTE)
25		Unassigned

Fig 4.1 : RS-232 interface connector pin assignments

PC/XT		NS	FUNCTION
1	----------------------	1	Protective Ground
2	<---------------------	2	Transmit Data
3	--------------------->	3	Receive Data
4	<---------------------	4	Request to Send
5	--------------------->	5	Clear to Send
6	--------------------->	6	Data Set Ready
7	--------------------->	7	Signal Ground
8	--------------------->	8	Data Carrier Detect
20	<---------------------	20	Data Terminal Ready

Fig 4.2 : Pin configuration for data transfer from NS to PC/XT

DTE (1 or 2) can now raise RTS to obtain control of the line. Data Terminal Ready (pin 20) plays a major role in the establishment ,maintenance of lines between any two communicating devices. Pin 20 is a control lead or signal used by the DTE to indicate that the modem should answer the call. The presence of DTR allows the connections and communication path to be established and maintained. If the far end computer doesn't have DTR on, the modem will not answer the call. However, for a connection to be maintained, the DTR signals should be present at both the ends. Data Set Ready (pin 6) from the DTE gives an indication regarding the status of the local data set. The 'ON' condition should not be interpreted as either an indication that a communication channel has been established to remote data station or the status of any remote station equipment. It simply indicates the connection between a local data communication equipment to a communication channel.

4.3 PRINCIPLES OF INTERFACING

Whenever cross connecting between devices, we should make sure that an output signal goes to an input signal, and vice versa. i.e. the transmitted data(output) is cross connected to the received data(input). Also, DTR(output) is connected to DSR (input). To permit a successful installation between devices, proper optioning with regard to speed, parity, character length, number of stop bits, mode of transmission, mode of operation, etc. should be executed. These terms are explained below to realize their importance in any interfacing set up.

Speed: The port speed of both devices, that are connected should be consistent to prevent data from being garbled. For example, if the maximum rate of operation for a printer is 300 bits per second (bps), the device sending data to the printer must also be set at 300 bps. The speed of transmission through a RS-232 port becomes extremely important when a printer is involved, due to the aspects of buffering and flow control. The speed of transmission from NS to Pc/XT has been set to 1200 bps.

Parity: The concept of keeping an odd/even number of bits in a character is known as parity. Parity could be even or odd depending on number of bits in a character. Besides keeping a fixed number of bits in a character, it also maintains the length of the character. Usually data is sent in terms of 1's and 0's (binary) between two computers. By counting the number of 1's, the receiving terminal does a parity check, thus ensuring the transfer of correct data. **If the parity is odd** and if the receiving computer detects even number of 1's then a parity error is said to have occurred. The causes of parity errors are numerous. Some of the more common causes being poor quality communication lines, power surges and poor interface connections. Any of these conditions may cause the flipping of one of the charcter's bits as it travels over the communication facility. However, to prevent the garbling of data it should be remembered that parity should be consistent at both ends. In some interfaces, parity is taken care by fixing the charecter length in terms of bits (say 8 bits). Such installations are said to have no parity, like the present set up.

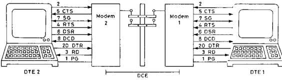


Fig.4.3 Interaction of RS-232 leads.

Character length : Another factor causing garbled data is the variation of character length. Again, the only remedy to avoid this, is to maintain the same character length at both the ends.

Number of stop bits : In technical terms, every data train has a start bit signifying the beginning of the data. The start bit informs the receiving computer that data bits will follow. A stop bit indicates the end of the character. Thus a character consists of a start bit, data bits and a stop bit. Generally stop bits could be 1 or 2. The present interface set up has 2 stop bits. However, for efficient and correct data transfer the same number of stop bits should be chosen for both ends of the configuration.

Mode : Generally, three modes of operation are available - simplex, half duplex and full duplex. Tracks allowing data transfer to go in one direction only are termed simplex. Half-duplex (HDX) or full duplex (FDX) deal with data flow in both directions. FDX offers simultaneous transfer of data in both directions unlike HDX, where only one direction could be chosen at a time. The present interfacing set up, in a way, could be called a simplex mode of operation as it offers transfer of data only in one direction, i.e. NS to PC/XT. Selection should be consistent, as this option could determine which RS-232 signals are generated or maintained by a port. We should choose the mode that suits the environment and application, and option both ends accordingly.

Mode of Transmission : Three choices are generally available for this option: asynchronous, synchronous and isochronous. The concept of enclosing a character with a start and stop bit is known as asynchronous transmission. The start bit indicates to the receiving computer, the time to start looking for the data and the stop bit indicates when the entire train of data had arrived. Thus the start and stop bits provide the timing. Each character is individually synchronized. In an asynchronous environment, we should be sure and check the number of stop bits option. Synchronous transmission involves transmission of data bits with a certain timing (speed limit). Protocols were developed to determine whether the data bits arrived correctly. Option one of the devices in a synchronous environment to provide the timing and check that the proper leads are present in the cables. The third choice is isochronous, which is a combination of the other two. Data in an asynchronous format are transmitted synchronously. We should ensure that both ends are optioned the same.

Polarity : Polarity has to do whether a signal is positive or negative. The importance of polarity is generally realized in the area of hardware flow control. However, the basic concept of maintaining the same polarity positive or negative at both ends is very important.

3.7.5 FILE TRANSFER BETWEEN NS AND PC/XT

To establish file transfer between NS and PC/XT, the serial port of the two systems has been exploited to provide a physical connection based on the RS-232-C standard. An important aspect of

RS-232 interfacing is whether the port is programmable or not. Programmability involves, as mentioned earlier, setting the baud rate, charcter length, parity, stop bits etc. This may be optainable through software or may require some hardware jumpering on the I/O board that provides the port.

The CP/M system on NS has a 8251 USART on its SIO board on slot 5 which is programmed for asynchronous transmission at a user designated baud rate. In the present interface set up, the COM1 of PC/XT was programmed at 1200 baud (bps), 8 bits , no parity and 2 stop bits through a program B.ASM on NS Advantage. This program sets the RS-232 line characteristic by programming the 8251 USART. This module programs the 8251 for the desired baud rate, number of bits/character etc. which exactly matches with parameters of the software (PC-VT) on the PC/XT side. The program B.ASM should be run once on the NS before every fie transfer session.

PC-VT is software emulator which makes the IBM PC function as a DEC VT102 or a VT 100 Video terminal. DEC, VT are trade marks of Digital Equipment Corporation. Here are a brief list of functions with reference to the present context, handled by PC-VT.

* Communication over an RS-232 port configured as either COM1 or COM2.
* Communication support for all the popular baud rate and data bit/parity bit combinations.
* Full receive parity checking can be enabled.

PC-VT requires an IBM PC with either the Monochrome or colour/Graphics card, an 80 column display, 64K of memory (i.e.

32K for Disc Operating System (DOS) plus 32K for PC-VT), one disc drive and one RS-232 port. The RS-232 port can be jumpered as COM1 or COM2. PC-VT supports the IBM fixed disc.

The files required to use PC-VT are PC-VT.EXE, PARAMS.DAT, DIALER.DAT, MACRO.DAT. These files should be there on the hard disc. The PC-VT program is started by typing PC-VT. PC-VT comes preinitialized to communicate at 1200 baud rate with 8 data bits, no parity and 1 stop bit. These parameters could be changed as per the user's requirement for the application.

As PC-VT starts up, it will initialize the COM port and establish the connection to the other computer/modem. The CRT will show the communication display. Lines 1 to 24 are used as emulator VT100/102 screen which has 24 lines. Line 25 is used as a status line. The word "ONLINE" and version number of PC-VT can be seen on the console at the lower left corner. This indicates that PC-VT is operating normally and is ready to communicate. In addition, the status line also shows the status of the attached modem/computer, to check if the cable is carrying all signals. These three icons are displayed:

- DSR - Data Set Ready - the modem/computer is ready to accept data.
- CD - Carrier Detect - the modem/computer has detected a carrier on the phone line.
- CTS - Clear to Send - the modem/computer has completed the establishment of a connection and is ready to send data to the host.

PC-VT has complete transmitting and receiving facilities. These facilities are accessed, using the CTRL-F4 and CTRL-F3 keys. These keys can be activated a communication display. For

receiving a file CTRL-F3 key is to be pressed, after which PC-VT will ask the name of the file to receive. Then the file name is typed on the screen giving details about the drive, path, file name with extension and the 'ENTER' key is pressed. "RCV" is displayed on the status line and any character sent by the host are copied to the PC disk file. The data/characters are displayed on the screen as and when they are transmitted. On completion of the transfer of data/text PC-VT is stopped by pressing CTRL-F8 and the file can be seen in the PC-VT directory.

The developed interface set up from NS to PC/XT offers two options which are listed and described below:

(i) Transfer of AET data files
(ii) Transfer of files other than AET

(i) Transfer of AET data files :

USERPRO is a 'skeleton program' provided by the AET (Acoustic Emission Technology) as a means for user's of AET 5000 to access and process AE data files stored on the disk. The program, as supplied by the system manufacturers, contains the necessary code for reading the AE data files.

On completion of cable connections between NS and PC/XT, after executing the program B.ASM on NS by simply typing 'B' followed by a carriage return, and after getting the PC-VT terminal ready to receive data, insert the disk containing the executable version of USERPRO (USERPRO.INT) and type "RUN USERPRO" and return. The program asks for AE data file to be read/transferred. After inserting the disk containing the AE data file, type the data file on NS and press return. At this stage three prompts appear on the screen:

(i) Print events?
(ii) Print statistical data?
(iii) Print rejected events?

The prompts have to be answered by typing Y or N. All the AE data files were transferred to PC/XT by answering Y,N,N in that order. Once, the prompts are answered, you would see AE data being displayed on the PC/XT monitor as the data is transmitted. Completion of AE data transfer session from NS is indicated by A>. Once, the transfer session is complete, the desired AE data file could be seen in PC-VT directory after coming out of the communication display. A typical transfer session from NS to PC/XT is shown in fig 3.11.

(ii) Transfer of non-AET files :

After carrying out the initial steps regarding (cable connection, B.ASM, PC-VT), as mentioned earlier, insert the disc containing the file to be transferred and along with wordstar files. Enter wordstar by typing WS on NS. Now print the file to be transferred through WS and you would see the file being transferred to PC/XT as it is printed. The transferred file could be seen in PC-VT directory.

Thus, the file transfer of AE data was accomplished from NS to PC/XT. Further, the data was transferred from PC/XT to DEC 1090 system, using the facility that is available at the Computer Centre. As a ready reference, general procedure for transfer of data from NS to PC/XT, PC/XT to DEC 1090 are listed at Appendix A and B respectively.

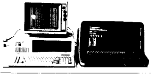


Fig. 4.4 Transfer of AE data from NS to PC/XT in progress.

CHAPTER 5

CLUSTER ANALYSIS

5.1 INTRODUCTION

One of the most primitive and common activities of man consists of sorting/classifying like things into categories. Classification is a process or act of assigning a new item or observation to its proper place in an established set of categories. Everyday we come across many types of problems, where in we predict the results, with reference to the past history. Be it credit analysis and insurance risks, screening people for disease during an epidemic or prediction of individuals for their success in an educational programme. In cluster analysis, little or nothing is known about the category structure. All that is available is a collection of observations whose category memberships are unknown. The operational objective in this case is to sort the observations into groups such that the degree of natural association is high among members of the same group and low between members of different groups. Cluster analysis has been employed in a great variety of applications, as an effective tool in scientific inquiry. The uses of cluster analysis in many fields of study could be broadly grouped into six major areas as listed below:

(i) Life Sciences (biology, zoology, ecology, paleontology)

(ii) Behavioral and Social Sciences (psychology, sociology, criminology, anthropology, linguistics,archaeology)

(iii) Earth Sciences (geology, geography, regional studies, soil sciences, remote sensing)

(iv) Medicine (psychiatry, cytology, clinical studies)

(v) Engineering Sciences (pattern recognition, artificial intelligence, systems science)

(vi) Information and Policy Sciences (operations research, information retrieval, political sciences, economics, market research)

It will be difficult perhaps to find a universally applicable definition of clustering, as any such definition is subject to variations depending on the investigators intuitive notion about what a cluster looks like, the criterion of clustering and the inherent groups present in the data. A widely used definition states clustering, as the process of finding homogeneous subsets in a set of data without using any a priori information regarding the class membership of individuals samples or events. An analogous definition is that clustering is the process of grouping various elements in a data set into subsets such that each point in a subset is more similar in some sense to other points in that subset than those in other subsets in the data set. Bow [33] defines clustering a nonsupervised classification of objects. It is the process of generating classes without any a priori knowledge of prototype classification. Thus for a given data set D of N events $\{x_1, x_2, \ldots, x_n\}$, to be classified into K clusters, the process of clustering can be formally stated as : to seek the clusters $c_1, c_2, \ldots, c_k$ such that every x_i, $i=1,2,\ldots,N$, fall into one of these clusters and no x_i falls in two region.

Mathematically,

$$c_1 \cup c_2 \;\ldots\; \cup c_k = D$$

$$c_i \cap c_j = \emptyset \quad \forall \; i = j \quad \text{where}$$

U and ∩ stands for union and intersection respectively.

5.2 STEPS OF CLUSTER ANALYSIS

A lot of discussion is devoted to all the choices leading to the final results. Anderberg [34] lists out atleast nine major steps in clustering a set of data, all of which can shape the outcome profoundly. These are briefly explained with reference to the present work.

choice of data units : When applying cluster analysis one should be clear regarding the objects for analysis. The objects could be persons, animals, opinions, commodities, or other such entities. In the present work, the objects for analysis was AE data obtained for Kevlar Epoxy composites, for five different configurations, as mentioned earlier.

There are two different situations of interest in the choice of data units. In the first instance, the available data is the complete object of the analysis and the purpose is to discover a classification scheme. The principal consideration here is to make sure that no important data units(events) are omitted. This was ensured when the AE data was analysed by the nonhierarchical methods of clustering. The second situation occurs more frequently. Here, the sample is a portion of a much larger population, which is the true object of interest. This was encountered when analysing AE data by hierarchical methods. These methods demand huge memory and massive computing time. Owing to these demands, only a portion of the available AE data sets could be analysed by hierarchical methods. Every second, third or fourth event was chosen depending on the size of the data sets.

choice of variables : It has the greatest influence on the ultimate results of cluster analysis. The data units cannot be merely arrayed for study. They must be described in terms of their characteristics, class memberships, traits and such other properties. Fiber reinforced composites exhibit a burst type of emission, characterized by the four AE parameters (PA,ED,RT,RDC). These are the variables chosen for the present study.

what to cluster : The discussion up to this point deal with the clustering of data units as described by variables. It is just as reasonable to cluster variables according to their mutual behaviour as manifested in the data sets. This view of thought is valid only when we have large number of variables. In our analysis of AE data, we have only four variables (infact, three in the present work, as RDC was suppressed due to its high correltation with ED. This would reduce CPU time to some extent). Therefore, the answer to the question is "AE data".

homogenizing variables : A common problem in real data is the lack of homogeneity among variables. In measuring association among variables different type of scales present difficult problems. In the present analysis, when dealing with nonhierarchical clustering methods, it was observed that the final clusters that were obtained were with reference to the dominant variable (dominant by virtue of its absolute value i.e. ED) To overcome this effect of the dominant variable,all the three variables (PA,ED,RT) were normalized between 0 and 100 using a linear interpolation. Thus, achieving the homogeneity among variable, which is necessary when clustering data units.

similarity measures : Most cluster analysis methods require a measure of similarity to be defined for every pairwise combination of the entities to be clustered. The similarity measure is usually given in numerical form to indicate the degree of natural association/resemblance between events in a data set. When clustering data units the proximity of individuals is usually is expressed as a distance. Two commonly used distances are the Euclidean distance (L2 metric) defined as :

$$d(X,Y) = [\ \Sigma_{i=1}^{m} \mid X_i - Y_i \mid^2\ ]^{1/2}$$

and the City block distance (L1 metric) defined as :

$$d(X,Y) = \Sigma_{i=1}^{m} \mid X_i - Y_i \mid \quad \text{where}$$

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix} \qquad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix}$$

m : number of events

Euclidean distance has been used as the similarity measure in the present analysis.

clustering criterion : Generally, one thinks of the term "cluster" as a set of objects which are all close together. But when it comes to finding clusters in real data, the term must bear a definite meaning. The choice of clustering criterion is tantamount to defining a cluster. It may not be possible to say just what a cluster is in abstract terms, but it can always be defined constructively through statement of the criterion and an implementing algorithm. In the present analysis, the clustering criterion for the nonhierarchical method is so formulated that

for a given matrix of I events on J variables, the events are allocated to N clusters in such a way that the within cluster sum of squares of deviations from their respective means is minimized.

algorithm and computer implementation : Even after choosing data units, variables, what to cluster, transformations, a similarity measure, and a criterion, there still remains the problem of actually generating a set of clusters. The choice of an algorithm should come after the other choices because it is merely the means of implementation. But very frequently, we find that all the choices are subordinated to the capabilities of an available computer/computer program with regard to its memory storage capabilities, handling number of events and variables and lastly the CPU time taken for the execution of programs. In the present study, DEC 1090 system was used for cluster analysis of AE data. It was found that the software developed for nonhierarchical clustering methods could be easily handled by the computer system whereas for hierarchical clustering methods, it could handle only a portion of the data sets for reasons that were mentioned earlier.

number of clusters : A substantial problem in performing a cluster analysis is deciding on the number of clusters in the data. The decision becomes all the more crucial, if you are dealing with nonhierarchical clustering methods. These methods do not exhibit the flexibility, that is offered by the hierarchical methods regarding the number of clusters. That is the reason, why one should be clear about the number of clusters that are needed. Hierarchical clustering methods give a

configuration for number of clusters from one (i.e.the entire data set) to the number of events (each cluster has only one event). Irrespective of the methods used, the number of clusters that were chosen for final analysis were limited to ten.

Interpretation of results : The finally produced clusters are summary descriptive statistics much like the mean and variance. For a given similarity measure, criterion and algorithm the production of clusters can be a straightforward mechanical procedure. But the whole focus of the clustering criterion and algorithm is to give a set of clusters that are well differentiated from each other. Even the analyst has got a role to play. The value of exploratory cluster analysis lies in suggested relationships and principles that were previously unnoticed. It is conditioned by three essential elements : the context of the problem, the analysts knowledge of the context and the analysts research objectives.

The context of the problem includes virtually everything that may influence the observation or interpretation of the data. The most pertinent aspect of context include the actual (versus intended) circumstances of data collection, related facts and theory. Also, the analysts knowledge and understanding of the problem context has long been recognised as an important element of all scientific inquiry. Even the analysts research objectives may change the face of entire investigation. Though a solution is sought with many objectives in mind, the ideal solution is the one which dominates, all other alternatives. Thus, context and objectives interact in cluster analysis, the later adjusting to

accommodate a new appreciation of the former.

5.3 CLUSTERING METHODS :

Clustering methods could be either hierarchical or nonhierarchical. In hierarchical clustering, the sequence of forming clusters proceeds in such a way that whenever two samples are put into one cluster at some stage, they remain together at all subsequent levels. In the nonhierarchical method, some initial seed points are chosen/obtained and the cluster memberships are altered with respect to certain criterion, before the final clusters are obtained.

5.3.1 NON-HIERARCHICAL CLUSTERING PROGRAM :

In the present work, emphasis was laid on the nonhierarchical clustering procedures for analysis of AE data obtained for composite specimens of five configurations and fiber rich specimens. The reasons for choosing this method for analysis is due to the advantages offered by this methods in comparison to hierarchical clustering method, as listed below :

1. Best suitable for large data sets for there is no need to either store the similarity matrix and the data set or to calculate the similarity matrix.
2. Demand less computing time .
3. Simple in operation .

Perhaps the only limitation of this method lies in the requirement of prior knowledge regarding the number of clusters needed. But once the user decides upon the number of clusters, these programs are very efficient, cost effective and need very less computational time.

A computer program was developed for the nonhierarchical clustering method. In this program, for a given matrix of I events on J variables, the events are allocated to N clusters in such a way that the within-cluster sum of squares of deviations from their respective means is minimized. The primary inputs are : matrix of observations, the number of clusters, and the initial cluster centres.

Selection of initial seed points or cluster centres is a very important part of any cluster analysis program. The seed points should be chosen as carefully as possible since this will have a direct influence on the running time of the program. Anderberg [34] lists some methods for generation of these initial seed points which are given below. These methods generate a set of N seed points which can be used as cluster nuclei around which the data set of I events can be grouped.

1. Choose the first N data units in the data set. If the initial configuration does not influence the ultimate outcome in any important way, then this method is the cheapest and simplest.
2. Label the data units from 1 to I and choose those labeled I/N,2I/N,..(N-1)I/N and I. This method is almost as simple as method 1 but tries to compensate for a natural tendency to arrange the events in the order of collection.
3. Subjectively choose any N data units from the data set.
4. Label the data units from 1 to I and choose the data units corresponding to N different random numbers in the range of 1 to I.

5. Take any desired partition of the data units in N mutually exclusive groups and compute the group centroids as seed points.

In all these above methods, though the selection of initial seed points is easy, they fail in one respect i.e. uniqueness. This is a very important property as far as initial seed points are concerned, especially so when the final clusters depend on the initial seed points. The method suggested by Ball and Hall [34] is by far the best method for selection of initial seed points. It offers a unique set of seed points for a certain threshold distance and certain number of clusters. The principle is : take the overall mean vector of the data set as the first seed point; select subsequent seed points by examining the data units in their input sequence and accept any data unit which is atleast some specified distance, say d, from all previously chosen seed points; continue choosing seed points until N seed points are accumulated or the data set is exhausted. This method is sufficiently simple that two or three values of the threshold distance could be tried if the first value gave too few seed points or examined too little of the data set. In addition to the versatility offered by this method, it was ensured that a large extent of data set has been examined before getting the required number of seed points. This method is less prone to give distorted or badly balanced clusters than methods listed above.

On the basis of the above principle,a computer program was developed for selection of initial seed points. The program is simple in operation and needs just four primary inputs : number of events, number of variables, number of seed points required

and the ofcourse, the data set. The threshold distance can be fed through the key board when prompted to do so. It was found that in most of the cases, a set of initial seed points were sought within 3 or 4 runs of the program. It is evident that for computational efficiency the threshold distance 'd' must be chosen as close as possible to the actual value to get the set of initial seed points within a lesser number of runs. A good initial guess is given by the formula :

$$d = a/N \left[\sum_{i=1}^{J} (R_i)^2 \right]^{1/2} \quad \text{------------ } 1$$

N : Number of clusters

J : Number of variables

R : Range of variable

Value of 'a' is chosen based on available information, regarding intercluster distances. In the absence of any such information, value of 'a' may be chosen arbitarily within the range $0<a<1$.

The AE data acquired on NS which was later transferred to PC/XT was in a very raw state for use by any cluster analysis program. A computer program was devaloped to format the raw data into a more compact form . A typical AE data output from NS and the compacted output of the same data are as shown in fig 5.1 and fig 5.2 respecyively.

Before executing the main program, a series of auxiliary programs have to be executed to obtain the primary inputs required by the main program. The purpose of the auxiliary programs are given below (see Appendix C for documentation):

1. To obtain the data set in the required form from raw

```
     TST SENSOR      DT1    DT2    DT3   DT4    DT5   DT6   DT7      ADC     ED  PA
ENG   RT  SLOPE    EVENT  TIME RG  LOC   PAR1   PAR2   PAR3

     # 1 12              1  ---    ---   ---    ---   ---   ---        9     39  44
 46   8      20  00:00:00.82  1     47    90     60   200
     # 1 12              9  ---    ---   ---    ---   ---   ---        3      8  39
 34   8      11  00:00:00.84  1     22    90     60   200
     # 1 12              7  ---    ---   ---    ---   ---   ---        8     33  41
 42  16       6  00:00:00.80  1     26    90     60   200
     # 1 12              1  ---    ---   ---    ---   ---   ---       50    323  75
 87  17     370  00:00:00.84  1     47    90     60   200
     # 1 21              8  ---    ---   ---    ---   ---   ---        2     16  32
 30  15       3  00:00:00.85  1     75    90     60   200
     # 1 12              1  ---    ---   ---    ---   ---   ---       13     51  44
 47   8      20  00:00:00.90  1     47    90     60   200
     # 1 12              0  ---    ---   ---    ---   ---   ---       56    241  77
 87   5    1413  00:00:00.91  1     50    90     60   200
     # 1 21              2  ---    ---   ---    ---   ---   ---       22     67  51
 55   7      51  00:00:00.91  1     57    90     60   200
     # 1 12              5  ---    ---   ---    ---   ---   ---        4     21  38
 36  18       5  00:00:00.94  1     34    90     60   200
     # 1 12              1  ---    ---   ---    ---   ---   ---       22     92  53
 59  17      25  00:00:00.96  1     47    90     60   200
     # 1 12              5  ---    ---   ---    ---   ---   ---        3     18  38
 37   1      79  00:00:00.98  1     34    90     60   200
     # 1 21              1  ---    ---   ---    ---   ---   ---       20     53  52
 55  11      35  00:00:00.99  1     54    90     60   200
     # 1 21              0  ---    ---   ---    ---   ---   ---        9     29  42
 41   8      13  00:00:01.07  1     75    90     60   200
     # 1 21              4  ---    ---   ---    ---   ---   ---       11     41  41
 43  10      11  00:00:01.06  1     53    90     60   200
     # 1 21              3  ---    ---   ---    ---   ---   ---        4     21  35
 34  17       3  00:00:01.12  1     60    90     60   200
     # 1 12              1  ---    ---   ---    ---   ---   ---       47    186  69
 78  17     165  00:00:01.13  1     47    90     60   200
     # 1 21              3  ---    ---   ---    ---   ---   ---       22     65  51
 55   8      44  00:00:01.17  1     60    90     60   200
     # 1 21              8  ---    ---   ---    ---   ---   ---       13     50  43
 46  11      13  00:00.01.29  1     75   150     80   200
     # 1 12              2  ---    ---   ---    ---   ---   ---        7     17  44
 42   8      20  00:00:01.37  1     44   150     80   200
     # 1 12             12  ---    ---   ---    ---   ---   ---        1      1  35
 21   1      56  00:00:01.45  1     13   150     80   200
     # 1 12              3  ---    ---   ---    ---   ---   ---        2     17  36
 34  16       4  00:00:01.48  1     41   150     80   200
     # 1 12              1  ---    ---   ---    ---   ---   ---        7     24  42
 42  21       5  00:00:01.51  1     47   150     80   200
     # 1 12              2  ---    ---   ---    ---   ---   ---       10     25  44
 44   8      28  00:00:01.57  1     44   150     80   200
```

Fig 5.1 Typical AE data output from Northstar Advantage

SAMPLE:RRS1

PA	ED	RDC	RT	ENG	SLP	PAR1	PAR2	PAR3	TST	SN	DT1	T1	T2	T3	RG	LOC
44	39	9	8	46	20	90	60	200	1	12	1	0	0	0.62	1	47
39	8	3	8	34	11	90	60	200	1	12	9	0	0	0.64	1	22
41	33	6	18	42	6	90	60	200	1	12	7	0	0	0.80	1	28
76	323	58	17	87	370	90	60	200	1	12	1	0	0	0.84	1	47
32	16	2	15	30	3	90	60	200	1	21	8	0	0	0.85	1	75
44	51	13	8	47	20	90	60	200	1	12	1	0	0	0.90	1	47
77	241	56	5	87	1413	90	60	200	1	12	0	0	0	0.91	1	50
51	67	22	7	55	51	90	60	200	1	21	2	0	0	0.91	1	57
39	21	4	18	38	5	90	60	200	1	12	5	0	0	0.94	1	34
53	92	22	17	59	26	90	60	200	1	12	1	0	0	0.96	1	47
38	18	3	1	37	79	90	60	200	1	12	5	0	0	0.98	1	34
52	53	20	11	55	36	90	60	200	1	21	1	0	0	0.99	1	54
40	29	9	8	41	13	90	60	200	1	21	8	0	0	1.07	1	75
41	41	11	10	43	11	90	60	200	1	21	4	0	0	1.08	1	63
36	21	4	17	34	3	90	60	200	1	21	3	0	0	1.12	1	60
69	195	47	17	78	165	90	60	200	1	12	1	0	0	1.13	1	47
51	65	22	8	55	44	90	60	200	1	21	3	0	0	1.17	1	60
43	50	13	11	46	13	150	80	200	1	21	8	0	0	1.29	1	75
44	17	7	8	42	20	150	80	200	1	12	2	0	0	1.37	1	44
35	1	1	1	21	56	150	80	200	1	12	12	0	0	1.45	1	13
36	17	2	16	34	4	150	80	200	1	12	3	0	0	1.48	1	41
42	24	7	21	42	6	150	80	200	1	12	1	0	0	1.51	1	47
44	25	10	8	44	20	150	80	200	1	12	2	0	0	1.57	1	44
42	25	9	8	42	16	150	80	200	1	12	2	0	0	1.61	1	44

Fig 5.2 : Typical ouput of AE data from FORMAT FOR

data obtained from AET 500 system.

2. To obtain the ranges of variables so that the initial threshold distance could be guessed.
3. To obtain scale factors for normalizing .
4. To obtain the initial seed points .
5. To remove the thinly populated clusters and outliers, which are considered due to noise .

The scale factors (say from 0 to 100) for the variables are calculated using a linear interpolation. The scale factors for all the variables are to be fed as data. The data set and the initial seed points are to be affected by the scale factors of the variables to achieve normalization (say from 0 to 100). The three variables are PA, ED, and RT. The ranges of a typical data set are taken to illustrate the calculation of scale factors and determining the threshold distance for generation of initial seed points.

For a typical data set the ranges are:

	MAX	MIN	RANGE
PA	77	23	54
ED	477	1	476
RT	95	1	94

As a sample calculation, scale factors for PA are obtained as shown:

$A = 100/54 = 1.85$ $\quad B = - 1.85 \times 23 = -42.55$

Similarly, scale factors are calculated for ED and RT to be fed as input.

The initial guess for threshold distance is calculated as shown:

$$d = 1/10\ [\ 54^2 + 478^2 + 94^2\ ]^{1/2} = 48.82$$

In a cluster analysis program the memberships of other clusters are affected by the presence of thinly populated clusters and outliers. Generally, these clusters consisted of high ED and RT events. After the initial execution of the main program such clusters were removed which are supposed to be due to noise. The number of clusters to be removed and their individual cluster number are to be fed as input for another program which eliminates the odd clusters and creates a modified data set. For this new data set, the scale factors and the inital seed points have to be determined before execution of the main program. Flow chart of main program is as shown in fig 5.3.

5.3.2 HIERARCHICAL CLUSTERING PROGRAM

Hierarchical methods are broadly divided into two types : agglomerative and divisive. In the agglomerative procedure, each of the events are treated as a cluster in beginning. The number of clusters are reduced by merging most similar pairs iteratively using a similarity measure till only one cluster remains or the desired number of clusters are obtained. If we call each of the events as a branch and the entire data set as a root, then, agglomerative methods can be described as clustering methods which build the tree from branches to the root. Less common are the divisive hierarchical methods, which begin at the root and work towards the branches. That is, the procedure begins by treating the entire data set as one cluster. This cluster is divided into two clusters using a similarity measure and the procedure is repeated till a suitable stopping criterion is

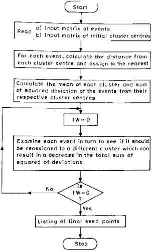


Fig. 5.3 Flow chart for Non-Hierarchical clustering program SCL·FOR.

satisfied.

The stored data approach has been used for the hierarchical clustering programs. This approach involves storing all the data in the central memory of the computer so that the entire body of numerical information is available about the problem. Then any clustering method with a similarity measure could be used because as original information is available on all data units. Though these programs demand massive computing time, they do offer some advantages in comparison to the nonhierarchical clustering methods such as 1) their suitability for small data sets and 2) flexibility in choosing the number of clusters.

Hierarchical method was specially chosen to analyse the fiber rich specimens for reasons mentioned above. A program was developed for this method based on the algorithm given in [34]. The main program makes use of five subroutines as shown in fig 5.4 to generate a interface file, to be used by the post processor program. Before executing the main program, the input data is to be modified as done in the nonhierarchical clustering methods. The input data contains the number of events, number of variables, in addition to the data set. The program structure and details of the subroutines are are as given in Appendix C.

A computer program was developed for post processing of merge data created in the interface file of the main computational program. This program requires one input i.e. number of clusters needed. The maximum and minimum number of clusters possible are the numbers of events in the data set and the whole data set respectively. Once the main program is

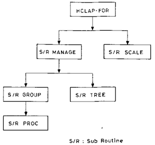


Fig. 5.4 Program structure for Hierarchical clustering program HCLAP·FOR.

executed for a given data, any number of clusters can be generated by the post processor program. This is definitely a positive point regarding the versatility of the program when compared to the nonhierarchical clustering programs. The cluster number, number of events in each cluster and the mean value of the variables are presented as an output.

At this stage, as in the case of nonhierarchical methods, the thinly populated clusters, outliers are removed. Another computer program was developed for this purpose which creates a modified data set. For the clustering of the modified data set, all steps as described in the previous paragaph are to be repeated.

In the present analysis of AE data for composite specimens and fiber rich specimens, 50 clusters were initially obtained from post processor program. At this stage, after removal of thinly populated clusters, the procedure was repeated as mentioned above and finally ten clusters were obtained.

For the analysis to be completed by hierarchical methods in a reasonable time (15 mts of CPU on DEC 1090) it was found that the data sets should be less than 1500 events for three variables. The fiber rich specimen data sets were small when compared to the composite specimen data sets. This facilitated execution of all events of fiber rich specimens. As the data sets of composite specimens were more than 2000 events, new data sets are formulated by choosing every second event if the total events are about 2500, every third if the total is 3000 events, every fourth if the total events are above 4000 and so on.

CHAPTER 6

RESULTS AND DISCUSSIONS

The AE data obtained by Rajan [35] from twelve specimens during his experimentation was analysed through cluster analysis to identify the damage mechanisms in Kevlar/Epoxy composites. Composites with following configuration were used:

a) Unidirectional laminates of $[0_{12}]$, $[90_{12}]$ and $[45_{12}]$ (two data sets per each configuration)

b) Crossply Laminates of $[0_3/90_3]_s$ and $[90_3/0_3]_s$ (two data sets per each configuration)

c) Specimens of fiber rich laminates (two data sets)

Table 6.1 gives an account of AE data sets with regard to their reference code, configuration and number of events.

6.1 SELECTION OF PARAMETERS

AET 5000 System characterizes signals from fiber reinforced composites by four parameters, namely peak amplitude (PA), event duration (ED), ring down counts (RDC) and rise time (RT). If an event with these four characteristic has to be identified with a damage mechanism, a method is to consider the position of event in a four dimensional space (4D). Events of the same characteristic when positioned in this 4D space would combine to form distinct clusters, identifying the different damage mechanisms.

Since, the concept of a 4D space is not easy to comprehend, a systematic cluster analysis is done. The selection of parameters was performed by studying the association of variables (correlation). A computer program was developed based on the

TABLE 6.1 : DETAILS OF AE DATA SETS

Configuration	Number of events	Reference Code
$[0_3/90_3]_s$	4292	A3.DAT
	3947	A2.DAT
$[90_3/0_3]_s$	2355	B3.DAT
	2588	B3.DAT
$[90_{12}]$	1084	C3.DAT
	2346	C2.DAT
$[0_{12}]$	2874	D3.DAT
	3077	D2.DAT
$[45_{12}]$	2852	E3.DAT
	2248	E2.DAT
FIBER RICH SPECIMENS	802	F1.DAT
	703	F2.DAT

TABLE 6.2 : CORRELATION BETWEEN AE VARIABLES

Sample : $[0_3/90_3]_s$

Data Set : A3.DAT

No. of variables : 4

	PA	ED	RDC	RT
PA		0.82	0.87	0.02
ED			0.92	0.07
RDC				0.07

algorithm given in [34] for determining the correlation between the variables. Typical correlation values between AE parameters calculated for a crossply sample are shown in table 6.2. It was observed that there was a strong correlation of 0.92 between ED and RDC. As such one parameter of the two (say ED) is sufficient for future analysis. Elimination of one parameter gave the following advantages :

1. It would ease the complexities in analysis with regard to computing time and number of iterations.
2. Absence of a fourth variable would facilitate easier analysis through 2D plots.

6.2 COMPARISON OF HIERARCHICAL AND NON - HIERARCHICAL METHODS :

AE data obtained for composite and fiber specimens was analysed by both hierarchical and non-hierarchical methods, with the software developed, as described in the previous chapter. In the non-hierarchical clustering method, the initial seed points were obtained by Ball and Hall method and the data sets were clustered for ten clusters. At this stage, the thinly populated clusters and the isolated events (outliers) were removed which are supposed to be due to noise. The modified data set, after removal of odd clusters was again clustered for ten clusters.

As mentioned before, hierarchical methods demand huge memory and massive computing time. Due to this inherent handicap of these methods, only a portion of the AE data sets were analysed by choosing every second, third and fourth event depending on the data size of the data set. The partial data sets were initially clustered for fifty clusters. At this stage, a total of 5% events

were removed which are attributed to the noise. A close study of the events in these clusters pertain to high ED and high RT events. Mechanical noise is characterized by high rise time or lower slope. After the removal, the modified data set was clustered for ten clusters.

In cluster analysis, the cluster number is just an identification number and does not give any information regarding the order of clustering. It is observed that ~~six~~ seven common clusters could be identified in table 6.3, which are obtained by two different clustering methods.

Let (X,Y) represent, cluster X obtained by non-hierarchical methods comparable with cluster Y obtained by hierarchical methods. The ~~six~~ seven clusters are (2,5), (3,8), (4,10), (5,6), (8,2) (7,1) and (9,7). The minor differences in the values of seed points obtained by these methods are due to the inherent way of clustering by each of these methods. In the non-hierarchical clustering methods, events are a assigned to the nearest cluster by calculating the distance from each cluster centre for each event. After the calculation of mean of each cluster and sum of squared deviation of events from their respective cluster centres, each event is examined in turn to see if it should be reassigned to a different cluster which can result in a decrease in the total sum of squared deviations. Thus, an event that is reassigned to particular cluster does not necessarily remain with the same cluster, till the final seed points are generated, unlike in hierarchical clustering, where the events once merged, remain in the same cluster till the final seed points are obtained. The final seed points obtained for all data sets of

composite specimens by non-hierarchical are given tables 6.4 A and 6.4 B.

Seven out of ten clusters being common, suggest a strong closeness of results obtained by both these methods. Similar behaviour was observed in the other data sets too. Thus, it was found that the final clusters obtained by both the methods yielded comparable results. This is quite evident, from results presented for a unidirectional sample (C3.DAT) in table 6.3. The satisfactory comparison of results by both these methods brings to light the fact that initial seed points generated by Ball and Hall method are good and balanced too.

TABLE 6.3 : COMPARISON OF FINAL SEED POINTS OBTAINED BY HIERARCHICAL AND NON-HIERARCHICAL METHODS FOR UNIDIRECTIONAL SAMPLE

Non-Hierarchical Methods					Hierarchical Methods				
CLNO.	EVENTS	PA	ED	RT	CLNO.	EVENTS	PA	ED	RT
1.	256	38.4	48.1	6.4	1.	111	34.5	39.7	13.9
2.	317	32.0	13.6	2.5	2.	38	33.1	36.6	25.4
3.	69	48.8	95.2	9.6	3.	183	40.9	86.2	9.6
4.	96	45.6	96.5	20.2	4.	183	37.0	36.3	4.7
5.	55	57.9	178.4	25.5	5.	281	31.7	11.7	2.6
6.	29	61.6	151.0	14.1	6.	53	56.5	181.4	25.1
7.	141	35.8	44.9	14.0	7.	27	36.6	60.5	40.6
8.	50	34.0	44.4	25.0	8.	50	52.9	111.1	11.7
9.	33	37.7	67.1	39.4	9.	27	39.8	77.2	24.7
10.	1	30.0	258.9	1.0	10.	70	46.5	102.6	20.1
Total	1023				Total	1047			

TABLE 6.4 A : FINAL SEED POINTS OF *2 AE DATA SETS

2A.DAT

CLNO.	EVENTS	PA	ED	RT
1	584	33.2	42.5	20.5
2	935	30.9	21.6	4.2
3	314	37.4	72.3	46.1
4	444	50.9	117.4	14.6
5	200	50.1	125.5	51.2
6	304	42.6	90.5	26.9
7	78	65.6	207.0	50.0
8	23	48.7	256.0	22.7
9	245	64.2	179.3	16.8
10	744	39.3	67.5	10.8

2B.DAT

CLNO.	EVENTS	PA	ED	RT
1	727	36.5	40.8	8.7
2	704	30.9	13.8	3.1
3	479	42.8	74.8	11.4
4	250	35.2	47.8	21.2
5	72	62.3	172.3	15.6
6	226	51.3	110.7	15.5
7	15	41.5	113.7	68.0
8	42	53.1	144.8	47.5
9	126	39.9	80.8	38.2
10	1	58.0	306.0	89.0

2C.DAT

CLNO.	EVENTS	PA	ED	RT
1	498	34.1	22.8	10.2
2	785	31.7	12.0	2.2
3	425	39.1	44.6	7.8
4	262	46.6	75.0	11.5
5	25	52.0	155.1	13.6
6	17	57.8	158.4	26.9
7	122	35.0	47.5	24.0
8	41	48.3	92.8	22.0
9	2	50.5	242.5	1.0
10	73	62.0	121.8	15.5

2D.DAT

CLNO.	EVENTS	PA	ED	RT
1	698	38.3	57.6	9.1
2	622	32.3	21.6	5.2
3	290	55.7	132.0	10.5
4	595	45.8	89.9	10.8
5	135	65.6	205.0	10.5
6	106	43.6	111.6	56.9
7	184	51.7	133.2	33.8
8	61	62.6	238.4	55.0
9	319	38.4	69.9	33.2
10	36	48.7	206.3	11.2

2E.DAT

CLNO.	EVENTS	PA	ED	RT
1	274	34.6	34.7	18.5
2	471	31.5	12.8	2.8
3	98	48.4	115.4	29.6
4	234	39.4	71.9	28.9
5	48	62.5	183.9	30.7
6	537	36.8	35.2	5.9
7	337	41.6	69.0	11.2
8	145	51.0	101.9	13.0
9	44	42.4	113.4	55.6
10	28	51.1	187.5	18.3

TABLE 6.4 B : FINAL SEED POINTS OF 3* AE DATA SETS

A3.DAT

CLNO.	EVENTS	PA	ED	RT
1	819	37.8	58.9	3.5
2	1033	30.3	16.9	2.2
3	545	46.3	97.6	5.4
4	400	56.2	137.6	6.2
5	175	68.0	215.1	6.6
6	261	36.8	69.6	39.3
7	413	33.2	43.1	20.5
8	158	48.9	124.8	30.8
9	306	40.9	83.1	20.7
10	89	62.0	195.1	32.4

B3.DAT

CLNO.	EVENTS	PA	ED	RT
1	535	34.9	29.3	5.2
2	498	30.3	10.5	2.4
3	445	38.9	54.8	8.8
4	244	45.2	76.6	9.3
5	81	35.3	55.8	30.3
6	74	43.4	87.0	25.6
7	207	32.8	29.1	13.5
8	115	54.8	116.7	13.0
9	25	47.4	154.3	12.0
10	44	39.0	92.7	54.7

C3.DAT

CLNO.	EVENTS	PA	ED	RT
1	256	38.4	48.1	6.4
2	317	32.0	13.6	2.5
3	69	48.8	95.2	9.6
4	96	45.6	96.5	20.2
5	55	57.9	178.4	25.5
6	29	61.6	151.0	14.1
7	141	35.8	44.9	14.0
8	50	34.0	44.4	25.0
9	33	37.7	67.1	39.4
10	1	30.0	258.9	1.0

D3.DAT

CLNO.	EVENTS	PA	ED	RT
1	495	41.1	72.4	6.1
2	558	33.2	29.1	4.2
3	465	49.6	104.8	10.7
4	283	60.1	158.1	10.8
5	40	68.0	259.1	38.8
6	412	38.2	61.0	19.6
7	204	49.1	118.1	26.8
8	200	39.3	81.0	42.8
9	70	53.2	157.2	48.8
10	125	70.8	259.1	8.8

E3.DAT

CLNO.	EVENTS	PA	ED	RT
1	612	35.7	34.3	6.7
2	155	42.5	85.0	25.9
3	641	31.4	11.7	2.3
4	493	41.0	61.9	12.0
5	221	48.3	99.4	15.0
6	27	58.6	179.6	29.7
7	378	35.0	35.4	16.5
8	107	35.5	50.1	32.0
9	91	55.5	143.8	19.4
10	52	45.5	116.8	42.7

6.3 PATTERN RECOGNITION ASPECT

Having two data sets for each configuration. First an attempt has been made to study the behaviour of samples of same configuration. After the clusters are obtained, for all AE data sets of composite specimens, two dimensional (2D) plots were generated between each pair of the variables: (PA,ED,RT) for each data set. Three combinations were possible for the 2D plots of

each data set which are PA Vs RT, PA Vs ED, ED Vs RT. These plots were later transferred to transparency sheets for further analysis. For each combination, two data sets of the same configuration were compared between each pair of parameters and those clusters which formed some definite patterns in each of the two plots, were cross checked with the other two combinations. For each configuration, the patterns formed by clusters between two different data sets were identified with regard to their parameters, disposition and behaviour. The patterns thus observed between the clusters were noted. Though the patterns were observed between the clusters, the absolute values of AE parameters were not same. the difference in the parameters could be attributed to specimen preparation, experimental condition and clamping of specimen-sensor contact. To study the effect of clamping, calibration experiments for specimens of different materials were conducted using different methods of contact. Clamping was accomplished using self adhesive tape or a C-clamp fixture. The results thus obtained are as given in table 6.b.

It could be seen that there is definite difference in the AE parameters depending on the method of clamping and the material Thus the shifts observed during the pattern recognition of clusters could be attributed the clamping of the sensor

Clusters involved in formation of patterns were averaged to give a single value of the pattern by taking the mean of parameters (PA, ED, RT). Thus, the two data sets for each configuration were reduced to a single data set by pattern recognition. Fig 6.6 to 6.11 represent the plots for A3.DAT and A2.DAT sets for $[0_3/90_3]_S$ involving all the three combinations.

The patterns that are formed between these sets are given in table 6.12.

TABLE 6.5 : EFFECT OF CLAMPING DEVICES

Material	Clamping device	PA	ED	RT	RDC
KEVLAR	Tape	74	44	5	12
	C-clamp	63	19	2	7
CARBON	Tape	72	132	12	31
	C-clamp	70	170	6	51

TABLE 6.12 : PATTERN FORMATION IN COMPOSITE SPECIMENS

A3.DAT	A2.DAT	B3.DAT	B2.DAT
1	10	1	1
9	6	2	2
6	3	3	3
7	1	10	7
8	5	5	9
10	7	7	4
2	2		
3	4		

C3.DAT	C2.DAT	D3.DAT	D2.DAT
2	2	1	1
1	3	2	2
4	8	3	4
5	6	4	3
6	10	5	8
7	1	10	5
8	7	9	6

E3.DAT	E2.DAT
1	6
4	7
5	8
7	1
6	5
3	2
8	4
2	3

Note: The numbers indicate cluster numbers of the data sets.

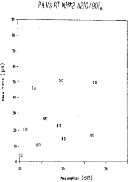


Fig.6.6 Plot of PA vs RT for A2.DAT

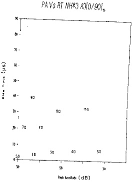


Fig.6.7 Plot of PA vs RT for A3.DAT.

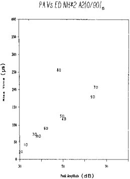


Fig. 6.8 Plot of PA vs ED for A2.DAT.

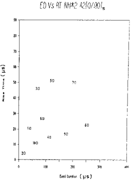


Fig. 6.9 Plot of ED vs RT for A2.DAT.

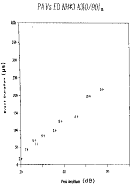


Fig. 6.10 Plot of PA vs ED for A3.DAT

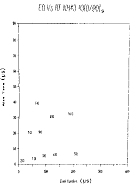


Fig. 6.11 Plot of ED vs RT for A3.DAT.

Due to their suitability to small data sets, hierarchical methods were chosen for clustering data of fiber rich specimens. After removing the events of thinly populated clusters at 50 cluster stage, the modified data was used to finally generate ten clusters. The same phenomena of pattern formation were observed in these specimens too. The two data sets of fiber rich specimens were reduced to one as per procedure given above. These specimens showed five clusters in common. Table 6.13 gives the formed patterns after merging.

TABLE 6.13 : PATTERN FORMATION IN FIBER RICH SPECIMENS

FR1.DAT	FR3.DAT	AE PARAMETERS OF #5 FR SPECIMENS		
		PA	ED	RT
4	3	48.2	57.6	13.6
10	7	55.6	87.8	16.7
8	8	71.1	189.2	20.1
9	10	64.4	131.5	14.0
5	4	41.2	32.5	11.6

After the generation of single data sets of composite and fiber rich specimens, three plots were again generated using all the three combinations of the variables. In our analysis these were referred as #5 data sets (say A5). Later, the plots were transferred onto the transparencies for the final analysis. At this stage, for each combination (say PA Vs RT) all the transparencies of the composite specimens were viewed by stacking them with the aim of identifying clusters containing clusters of different configuration, but characterizing the same damage

mechanism. A group is selected if only if it is present in all the three combinations. Three such groups were identified whose details are given in table 6.14.

TABLE 6.14 : CLUSTERS IDENTIFYING DAMAGE MECHANISM

Group	Membership	PA	ED	RT	EVENTS	TOTAL EVENTS
	A7	30.6	19.3	3.2	1968	7229
	B2	30.6	12.2	2.8	1202	4111
I	C1	31.9	12.8	2.4	1102	2905
	D2	32.8	25.4	4.7	1180	4543
	E6	31.4	12.3	2.6	1112	4778
	A1	38.6	63.2	7.2	1563	7229
	B3	40.9	64.8	10.1	924	4111
II	C2	38.8	46.4	7.1	881	2905
	D1	39.7	65.0	7.6	1193	4543
	E2	41.3	65.5	11.6	830	4778
	A4	33.2	42.8	20.5	997	7229
III	B6	34.0	38.5	17.4	207	4111
	C8	34.5	45.9	24.5	639	2905
	E4	34.8	35.1	17.5	652	4778

The minor differences in AE parameters in the cluster groups can be attributed to specimen preparation, experimental conditions and sensor-specimen contact. Differences were observed in AE parameters, when the sensor was attached to the specimen surface with the couplant (SO6) along with self adhesive tape and when bound by a C-clamp made for this purpose.

In a nut shell, the ranges of the three parameters for each group are given in table 6.15.

TABLE 6.15 : RANGES OF AE PARAMETERS FOR DAMAGE MECHANISMS

GROUP	PA		ED		RT	
	lower	upper	lower	upper	lower	upper
I	30.0	33.6	0.0	32.3	0.0	10.7
II	35.6	42.8	41.9	74.2	0.0	12.9
III	30.0	35.1	32.3	64.5	14.3	28.6

These three groups have lower PA denoting that they belong to matrix related damages such as matrix cracking, fiber-matrix debond, delamination etc. The ranges of the parameters are also included in the above table, which when given as an input, while postprocessing the AE data or as real time parameters, can identify these three damages.

In the analysis of fiber rich specimens, it was found that it is not only fiber break that is dominant as one would expect. Several other damage mechanisms were also found to occur, identification of which needs finer experimentation.

Thus, the cluster groups were identified for certain damage mechanisms in Kevlar epoxy composites. However, pin pointing of the failure process needs, a further in-depth study and model experiments.

CHAPTER 7

CONCLUSIONS

7.1 CONCLUSIONS

Cluster analysis of AE data, obtained from kevlar/epoxy composites was carried out with an aim to identify the various damage mechanisms. The data was acquired using the AE technique. The analysis was conducted on AE data obtained for unidirectional laminates of 0°, 90°, 45° - configurations, cross-ply laminates of $[0_3/90_3]_s$ and $[90_3/0_3]_s$ and fiber rich specimens. Based on observations made during the analysis, the following conclusions are drawn:

1. Cluster analysis offers a better insight to damage identification in composites when compared to statistical analysis.
2. Three groups (two five member groups and one four member group) were identified through cluster analysis of AE data representing certain damage mechanisms. Matrix failures are generally associated with low PA and ED events. Hence, the presence of low PA and ED events in Group 1 may be associated to matrix failure.
3. Though groups 2 and 3 definitely identify certain damage mechanisms, pin pointing of the failure process needs a further in-depth study and controlled model experiments.

7.2 SCOPE FOR FUTURE WORK

The present work has laid emphasis on development of a methodology to identify damage mechanisms in kevlar epoxy composites through cluster analysis of AE data. In order to exactly identify the damage mechanisms like delamination, the future scope of work lies in development of specially tailored experiments, so as to obtain the exact AE characteristics of various failure processes. It would be interesting to find the damage progression with respect to load to attain information from the onset of failure to fracture.

More exhaustive pattern recognition methods could be employed for feature extraction and damage identification. The most sophisticated level of damage identification could be an on-line technique, incorporating failure characterization by just looking at the AE parameters.

REFERENCES

1. J. Czochralski, "Die Metallographie des Zinns und die Theories der Formanderung Bildsamer Metalle", pp 331-593, Metall und Erz, Sep 1916 (in German)

2. F. Forster and E.Scheil, "Akustische Untersuchung der Bildung von Martensitnadeln (Acoustic Study of the Formation of Martensite Needles)", pp 245-247, Zeitschrift fur metallkunde, Sep 1936 (in German).

3. W.P. Mason, H.J.McSkimin and W.Shockley, "Ultrasonic Observation of Twinning in Tin", pp 1213-1214, Physical Review, May 1948.

4. D.J.Millard, "Twinning in Single Crystals of Cadmium", Ph.D. Thesis, University of Bristol, Bristol, England, 1950

5. J.Kaiser, "Untersuchungen uber das Auftreten Gerauschen beim Zugversuch (An Investigation into the Occurance of Noises in Tensile Tests)", Ph.D. Thesis, Technische Hochschule Munchen, Munich, Germany, 1950 (in German).

6. H.L.Dunegan and D.O. Harris, "Acoustic Emission: A New Nondestructive Testing Tool", Lawrence Radiation Laboratory, Livermore, California, pp 203-238, Symposium on the NDT of welds and Material Joining, ASTM, Evanston, Illinois, Jan 1968.

7. A.A.Pollock, "Acoustic Emission Methods of NDT", pp 85-89, British Journal of Non-Destructive Testing, May, 1971.

8. R.T.Anderson and T.J DeLacy, "Nondestructive Testing of Advanced Composites", pp 88-92, Metal Progress, Aug. 1972

9. M.A.Hamstad, "Acoustic Emission Uses in Research and Development of Composites Materials", Lawrence Livermore Laboratory, livermore, California, Nov 1974

10. J.M.Carlyle, "Acoustic Emission in Fiber Reinforced Composites", M.Tech. Thesis, Brunel University, Uxbridge, England, Sep 1974

11. F.J.Guild, D.Walton, R.P.Adams and D Short, "The Application of Acoustic Emission to Fiber Reinforced Composite Materials", pp 173-179, Composites, July 1976

12. M.Arrington, "Acoustic Emission from Real and Model Composites", pp 3-1 - 3-5, Newsletter on the Non-Destructive Testing of Composite Materials and Adhessive Bonded Joints, Issue No.2, Royal Aircraft Establishment, Farnborough, Hants, England, July 1973.

13. H.Ahlborn, J.Becht and J.Eisenblatter, Anwendung der Schallemissionsanalyse bei Untersuchungen uber die Risszahigkeit von Faserverbundwerkstoffen (Applications of Acoustic Emission Analysis for Studying Fracture Toughness of Composites)", Battelle-institut eV, Frankfurt am Main, Germany, 1973.

14. A.R.Bunsell, M.Fuwa and B.Harris, "Evaluation Studies of Acoustic Emission",pp 7-1 - 7-7, Newsletter on the Non Destructive Testing of Composite Materials and Adhessive

Bonded Joints, Issue No.3, Royal Aircraft Establishment, Farnborough, Hants, England, June 1974.

15. C.R.Brown, "Acoustic Emission from GRP", pp [illegible] [illegible], Newsletter on the Non-Destructive Testing of Composite Materials and Adhesive Bonded Joints, Issue No.4, Royal Aircraft Establishment, Farnborough, Hants, England, July 1975.

16. T.T.Chiao, M.A.Hamstad, M.A.Marcos, and J.E.Hanafee, "Filament - Wound Kevlar 49 / Epoxy Pressure Vessels", Lawrence Livermore Laboratory, Livermore, California, Nov 1973.

17. M.A.Hamstad and R.G.Patterson, "Considerations for Acoustic Emission Monitoring of Spherical Kevlar - Epoxy Composite Pressure Vessels", Lawrence Livermore Laboratory, Livermore, California, May 1977.

18. A.R.Bunsell, "The Monitoring of Damage in Carbon Fibre Composite Structures by Acoustic Emission", Composite Structures (2), 1983.

19. Richard.L.Moore and Marvin A.Hamstad, " Acoustic Emission from Single and Multiple Kevlar-49 Filament Breaks", Journal of Composite Materials, Vol. 20, January 1986.

20. S.Ghaffari and J.Awerbuch, "Monitoring Acoustic Emission in Impact-Damaged Composites", pp 120-125, Proceedings of the Second International Symposium on Acoustic Emission from

Reinforced Composites, July 1986.

21. J.Awerbuch and S.Ghaffari, "Monitoring Progression of Matrix Splitting During Fatigue Loading through Acoustic Emission in Notched Graphite / Epoxy Composites", pp 51-58, Proceedings of the Second International Symposium on Acoustic Emission from Reinforced Composites, July 1986.

22. J.Awerbuch and S.Ghaffari, "Tracking Progression of Matrix Splitting During Static Loading through Acoustic Emission in Notched Unidirectional Graphite / Epoxy Composites"

23. M.Clerico, E.Ceccarelli and G.Magistrali, "On Detection of Damage in Advanced Composites", pp 1705-1714, Non-destructive Testing, edited by J.M.Farley, R.W.Nichols, 1986.

24. J.Cohen and J.Awerbuch, "Monitoring Delamination Progression in Graphite / Epoxy Composite through Acoustic Emission During Fatigue Loading".

25. J.Cohen and J.Awerbuch, "Tracking Progression of Delamination in Model Graphite / Epoxy Composites through Acoustic Emission During Quasi-Static Loading", National Center for Composite Materials Research", University of Illinois, Oct 1987.

26. M. Tasnon, C.Roy and A.Maslouhi, "Damage Monitoring and Characterization in Carbon / Epoxy Composites", pp 291-299, Analytical and Testing Methodologies for Design with Advanced Materials, North-Holland, 1988.

27. J.Awerbuch and J.G.Bakuckas, "On the applicability of Acoustic Emission in Metal Matrix Composites", Metal Matrix Composites: Testing, Analysis and Failure Modes, ASTM, Philadelphia, 1989

28. J.H.Williams, Jr. and D.M.Egan, "Acoustic Emission Spectral Analysis of Fiber Composites Failure Mechanisms", pp45-47, Materials Evaluation, January 1979.

29. T.Prasanna Kumar, G.Jothinathan and Dr O.Prabhakaran, "The Acoustic Emission Technique (AET) as a Material Testing Tool, Indian Institute of Technology, Madras, 1984

30. M.G.Philips, F.J.Ackerman and B.Harris, "A Study of Acoustic Emission Amplitude Distribution From Composites, Using Simple Statistical Techique", pp 440-446 in Ultrasonics International 81 - conference proceedings, IPC Science and Technology press limited, Surrey, England, 1981

31. M.Abdul Majeed, "Unsupervised Pattern Classification of Acoustic Emission Signals", M.S.Thesis, Indian Institute of Science, Bangalore, February 1987.

32. M.D.Seyer, "RS-232 Made Easy", Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1984.

33. S.T.Bow, "Pattern Recognition", Marcel Dekker, Inc. New York, 1984.

34. M.R.Anderberg, "Cluster Analysis for Applications", Academic Press, New York and London, 1973.

35. D.S.Rajan, "Investigation of Damage in Kevlar/Epoxy composites using Acoustic Emission Technique", M Tech. Thesis, Indian Institute of Technology, Kanpur, May 1988.

APPENDIX A

PROCEDURE TO TRANSFER AE DATA FROM NORTHSTAR TO PC/XT

ON PC/XT:

1. CD \PC-VT
2. PC-VT
3. Press any key
4. CTRL F3 (F3) : It will ask you which file to be received. Give the file name and press enter twice. PC VT is ready to receive the data.

ON NORTH STAR:

1. Load the system disc containing the file B.COM and CARRIAGE RETURN <CR>.
2. System displays A>. Execute B.ASM by typing B, followed by <CR>

TO TRANSFER FILES OTHER THAN AET FILES

3. Load the disc containing the file to be transmitted and WORDSTAR files, followed by <CR>.
4. Enter WS by typing WS, followed by <CR>.
5. Print the file to be transmitted through WS. The file appears the PC/XT screen as it prints.

TO TRANSFER AE DATA FILES

6. Load the disc containing the file USERPRO.INT.
7. Execute the above program by typing "RUN USERPRO", followed by <CR>.

8. USERPRG asks you for the data file to be transferred

9. Load the disc containing the AE data file and type the name of the file with extension, followed by <CR>. Note that all AE data files have a common extension (.101)

10. Three prompts appear on the screen at this stage:

    i) Print events? ii) Print statistical data? iii) Print rejected events?. Answer them "Y" or "N" as per your requirement. The AE data files were transferred by answering Y, N, N in that order. System will come to A> on completion of data transfer. Infact, this is the indication that the data transfer is completed

ON PC/XT:

5. Once the transfer of file is completed, terminate PC-VT by pressing CTRL - F8 (^F8). You can see the file in PC-VT dir.

Procedure is same for subsequent transfers, but you don't have to run program B. It is to be run once for a transfer session. Program B programs the COM1 port of PC/XT to receive the data at 1200 bps(bits/sec), 8 bits, no parity, 2 stop bits.

PROCEDURE TO ASSEMBLE B.ASM

1) ASM B <CR> ;This creates object file B.OBJ

2) LOAD B <CR> ;This creates executable file B.COM

APPENDIX B

PROCEDURE TO TRANSFER AE DATA FROM PC/XT TO DEC 1090 SYSTEM

1. CD\PC VT
2. PC-VT
3. Press any Key
4. CTRL - C (^C) : The terminal comes to dot (.) mode.
5. Login
6. Copy DEC10.DAT = tty: <CR>
7. CTRL -F4 (^F4) : It will ask for the file to be transferred.
8. Give the file name and <CR>. For eg: a:A3.DAT <CR>
9. "Transmission Started" will appear on the screen, indicating beginning of transfer of data.
10. Once the end of data file is reached, it will ask "Should ^Z be transmitted ? Answer Y or N to this prompt.

Steps 6 to 10 are to be repeated for subsequent transfer of data files. On completion of file transfer, the job is killed by typing KJ, followed by a carriage return <CR>. PC-VT is terminated by pressing CTRL - F8 (^F8)

APPENDIX C

DOCUMENTATION FOR CLUSTER ANALYSIS PROGRAMS

Non-hierarchical Clustering Program:

Before executing the main program SCL.FOR, a series of auxiliary programs have to be executed to obtain the primary inputs required by the main program. The auxiliary programs and their functions are given below :

CUT.FOR	To obtain the data set from FORMAT.FOR as required by SCL.FOR .
SEED.FOR	To obtain the ranges of variables so that the initial threshold distance could be guessed for BALL.FOR .
	To obtain scale factors for normalizing between 0 to 100.
BALL.FOR	To obtain the initial seed points .
REMOVE.FOR	To remove the thinly populated clusters and outliers.

Hierarchical Clustering Program

HCLAP.FOR is the main program. It makes use of five subroutines as shown to generate a file INFACE.IN, to be used by the post processor program POSTAN.FOR. Before executing the main program, the input file RAW.DAT is to be generated through program CUT.FOR. The input file contains the number of events, number of variables, in addition to the data set. The program structure is as given below:

HCLAP.FOR	It establishes the dimension of various arrays, checks for sufficient storage requirements. Reads the input file RAW.DAT. Calls subroutine SCALE on users decision. Calls subroutine MANAGE.

MANAGE It controls the execution for a hierarchical clustering job based on an input data vector read in the main program. Calls subroutine GROUP and TREE.

GROUP Calls subroutine PROC iteratively.

PROC It computes the total within group sum of squared deviations.

TREE It creates the necessary files for post processing. The tree drawing part has been suppressed for it occupies huge memory space.

SCALE Scales the variables from 0 - 100 according to users decision.

Once the input file RAW.DAT is ready, HCLAP.FOR can be executed straight without any modification, if the total number of events are less than 2000. As the CPU time is more for large data sets, it is advisable to restrict the number of events to 1500 as suggested by the CPU chart given below. In this case, there is no need for any modifications in the program.

Number of Events	CPU time in minutes
1000	2.5
2500	15.0

FOSTAN.FOR does the post processing of merge data created in file INFACK.IN of the main computational program HCLAP.FOR. This program just requires one input i.e. number of clusters needed. The maximum and minimum number of clusters possible are the numbers of events in the data set and the whole data set respectively. Once the main program is executed for a given data, any number of clusters can be generated by just executing FOSTAN.FOR. The cluster number, number of events in each cluster and the mean value of the variables are given in the output file MEAN.OUT.

At this stage, as in the case of nonhierarchical methods, the thinly populated clusters, outliers are removed. This is

accomplished by the program REMOV.FOR which creates a modified data set TRAW.DAT. For the clustering of the modified data set, all steps from BCLAF.FOR to POSTAN.FOR are to be followed after copying TRAW.DAT into RAW.DAT.

ME - 1989 - M - KIM - DAM